VOL. 32, NO. 5, May-June, 1959 PRICE 65¢

MATHEMATICS magazine

Address editorial correspondence to Glenn James, special papers to the editors of the departments for which they are intended, and general papers to a member of the executive committee.

Manuscripts should be typed on 8½″× 11″ paper, double-spaced with 1″ margins. We prefer that, in technical papers, the usual introduction be preceded by a *Foreword* which states in simplest terms what the paper is about. Authors need to keep duplicate copies of their papers.

The Mathematics Magazine is published at Pacoima, California by the managing editor, bi-monthly except July-August. Ordinary subscriptions are: 1 year \$3.00; 2 years \$5.75; 3 years \$8.50; 4 years \$11.00; 5 years \$13.00. Sponsoring subscriptions are \$10.00; single copies 65¢; reprints, bound 1¢ per page plus 10¢ each, provided your order is placed before your article goes to press.

Subscriptions and other business correspondence should be sent to Inez James, 14068 Van Nuys Blvd., Pacoima, California.

Entered as second-class matter, March 23, 1948, at the Post Office, Pacoima, California, under act of Congress of March 8, 1876.

MATHEMATICS MAGAZINE

VOL. 32, NO. 5, MAY-JUNE, 1959

CONTENTS

Arithmetic

- Do you know arithmetic?
- I am a master of it.
- How do you divide four dollars among three people?
- Give two dollars to each of the first two, and let the third one wait until two more dollars are provided.

DIFFERENCE EQUATIONS AND THEIR APPLICATIONS

Louis A. Pipes

I. *Introduction.*

During the last fifteen years there has been a great deal of activity in the applications of difference equations in the solution of many problems that arise in statistics, science, and engineering.

The development of high-speed digital computing machinery has motivated the use of difference equations as approximations to ordinary and partial differential equations. [See bibliography, 1, 2, 3]. The use of difference equations for solving partial differential equations was discussed in 1928 in a celebrated paper by Courant, Friedrichs, and Lewy. [4]. However, the methods presented in this paper were not put into practical use until about 1943 as a result of the stimulus of war-time technology and with the assistance of the first digital automatic computers. Problems involving time-dependent fluid flows, neutron diffusion and transport, radiation flow, thermo-neuclar reactions, and problems involving the solution of several simultaneous partial differential equations are being solved by the use of difference equations throughout the country.

Besides the use of difference equations as approximations to ordinary and partial differential equations, they afford a powerful method for the analysis of electrical, mechanical, thermal, and other systems in which there is a recurrence of identical sections. The study of the behavior of electric-wave filters, multistage amplifiers, magnetic amplifiers, insulator strings, continuous beams of equal span, crankshafts of multicylinder engines, acoustical filters, etc., is greatly facilitated by the use of difference equations. The usual methods for solving such systems are generally very lengthy when the number of elements involved is large. The use of difference equations greatly reduces the complexity and labor in problems of this type.

II. *Nature of the Calculus of Finite Differences.*

As is well-known, the most important concept of mathematical analysis is that of a function. If to a given value of x a certain value of y corresponds, we say that y is a function of the independent variable x and write symbolically $y(x)$: Two types of functions occur in the applications of mathematical analysis to physical problems; first there are functions in which the variable x may take every possible value in a given interval, that is, the variable x is continuous. Functions of this type are studied in the ordinary *Differential and Integral Calculus*.

A second type of function exists in which the variable x takes only the given values $x_1, x_2, x_3, \cdots x_n$. In this case the variable x is discontinuous. The methods of the ordinary calculus are not in general applicable to such functions. *The Calculus of Finite Differences* deals especially with functions of the second type, but it may be applied to functions of the first type as well. The origin of the Calculus of Finite Differences is usually ascribed to Brooks Taylor who published his, "Methodus Incrementorum" in 1717. He pointed out that the relations between the terms of arithmetic and geometric progressions, which were known for a long time represented the simplest examples of difference equations.

There are many analogies between the Differential and Integral Calculus and the Calculus of Finite Differences. For example, in the Calculus of Finite Differences, the first difference, denoted by $\Delta(h)F(x)$ is defined by,

$$\Delta(h)F(x) = F(x+h) - F(x) \tag{2.1}$$

In the Differential Calculus, the first derivative of a function $F(x)$ is defined by the equation,

$$DF(x) = \operatorname*{Limit}_{h\to 0} \frac{[F(x+h)-F(x)]}{h} = \operatorname*{Limit}_{h\to 0} \frac{[\Delta(h)F(x)]}{h} \tag{2.2}$$

It can be shown that the nth. derivative of $F(x)$ may be expressed in the form,

$$D^nF(x) = \operatorname*{Limit}_{h\to 0} \frac{[\Delta^n(h)F(x)]}{h} \tag{2.3}$$

The operator $\Delta(h)$ is called the *forward difference operator*. When there cannot be any misunderstanding about the increment h, the difference operator $\Delta(h)$ is merely written as Δ.

Basic Operators used in the Calculus of Finite Differences.

The following fundamental operators are used widely in the Calculus of Finite Differences.

1. *The Displacement Operator E.*

$$EF(x) = F(x+h) \tag{2.4}$$

2. *The Forward Difference Operator Δ.*

$$\Delta F(x) = F(x+h) - F(x) \tag{2.5}$$

3. *The Backward Difference Operator ∇.*

$$\nabla F(x) = F(x) - F(x-h) \tag{2.6}$$

4. *The Central Difference Operator* δ.

(2.7) $$\delta F(x) = F(x+\frac{h}{2}) - F(x-\frac{h}{2})$$

5. *The Averaging Operator* μ.

(2.8) $$\mu F(x) = \frac{1}{2}[F(x+\frac{h}{2}) + F(x-\frac{h}{2})]$$

6. *The Symbolic Representation of Taylor's Expansion.*

(2.9) $$e^{hD}F(x) = F(x+h), \qquad D = \frac{d}{dx}$$

Equivalent Operators.

By using the above fundamental definitions of the basic operators used in the calculus of finite differences, the following equivalent operators may be easily deduced.

(2.10) $$E = e^{hD}$$

(2.11) $$\Delta = (E-1)$$

(2.12) $$\nabla = (1-E^{-1})$$

(2.13) $$\delta = E^{1/2} - E^{-1/2}$$

(2.14) $$\mu = \frac{1}{2}\delta = \frac{1}{2}(E^{1/2} + E^{-1/2})$$

It can be shown [5, 6, 7, 8] that except for a few restrictions, these operators can be manipulated according to the laws of ordinary algebra, and many relations between the basic operators may be obtained.

For example, we may write,

(2.15) $$D = \frac{\ln(1+\Delta)}{h} \quad \text{by the use of (2.10)}$$

(2.16) $$\nabla = (1-E^{-1}) = (1-e^{-hD})$$

(2.17) $$\delta = 2\sinh(hD/2)$$

(2.18) $$\mu = \cosh(hD/2)$$

An interesting and useful exposition of the symbolic calculus of operators will be found in a paper by Bickley [9]. By the use of these operators, and their relations between them, several finite difference formulae for interpolation, differentiation, integration, scanning, etc. may be obtained. These formulae are of great importance in numerical analysis.

For example, if the logarithm in (2.15) is expanded in powers of Δ,

we obtain,

$$D = \frac{\ln(1+\Delta)}{h} = \frac{1}{h}(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \cdots) \tag{2.19}$$

This expression may be used to compute the derivative of a function in terms of its forward differences.

The Scanning Operator. $S(h)$

In the theory of scanning, if we are given a function $F(z)$, it is required to obtain the function $\phi(x)$ defined by,

$$\phi(x) = \int_{(x-h/2)}^{(x+h/2)} F(z)dz \tag{2.20}$$

If (2.20) is differentiated with respect to x, the result is,

$$\begin{aligned} D\phi(x) &= F(x+\frac{h}{2}) - F(x-\frac{h}{2}) \\ &= \delta F(x) \end{aligned} \tag{2.21}$$

Hence,

$$\phi(x) = \frac{\delta}{D}F(x) = S(h)F(x) \tag{2.22}$$

Where $S(h)$ is the *scanning operator* defined by,

$$S(h) = \frac{\delta}{D} = \frac{2\sinh(hD/2)}{D} = \frac{h\sinh(hD/2)}{(hD/2)} \tag{2.23}$$

If $\phi(x)$ is known and it is desired to determine $F(x)$, we may write,

$$F(x) = S^{-1}(h)\phi(x) \tag{2.24}$$

In (2.24), $S^{-1}(h)$ is the *inverse scanning operator.*

III. *Linear Difference Equations with Constant Coefficients.*

The most important difference equations that arise in applications to dynamics, electrical circuits, structural problems, etc., are linear difference equations with constant coefficients.

These equations are of the type,

$$(a_nE^n + a_{n-1}E^{n-1} + \cdots + a_1E + a_0)y(x) = V(x) \tag{3.1}$$

In the usual applications of difference equations, $h = 1$, so that the operator E is defined by,

$$E^k y(x) = y(x+k) \tag{3.2}$$

Difference equations are classified in the same manner as differential equations. The equation (3.1) is called a linear difference equation of order n with constant coefficients, provided that the coefficients a_k are constants. If $V(x) = 0$, then the equation is said to be *homogeneous*; if not, it is said to be *non-homogeneous*.

Let,

$$P(E) = (a_n E^n + a_{n-1} E^{n-1} + \cdots + a_1 E + a_0) \tag{3.3}$$

The general solution of (3.1) may be written in the form,

$$y = y_c(x) + y_p(x) \tag{3.4}$$

If this is substituted into (3.1), the result is,

$$P(E)(y_c + y_p) = 0 \tag{3.5}$$

Let,

$$P(E)y_c(x) = 0, \tag{3.6}$$

and

$$P(E)y_p(x) = V(x) \tag{3.7}$$

By analogy to the solution of linear differential equations with constant coefficients, y_0 is called the *complementary function,* and y_p the *particular integral* of the differential equation (3.1).

Determination of the Complementary Function.

To determine the complementary function $y_c(x)$ we assume a tentative solution of the form,

$$y_c(x) = Aq^x \tag{3.8}$$

where A is an *arbitrary constant* and q a number to be determined so that (3.8) is a solution of (3.6). It is noticed that if the operator E^k is applied to (3.8) the result is,

$$E^k y_c(x) = E^k Aq^x = q^k Aq^x = q^k y_c(x), \tag{3.9}$$

It is therefore evident that if (3.8) is substituted into (3.6), the result may be written in the form,

$$P(E)y_c(x) = P(q)y_c(x) = 0 \tag{3.10}$$

If the trivial solution $y_c(x) = 0$ is excluded, then we must have,

$$(3.11) \qquad P(q) = 0$$

This is the *characteristic equation* of the difference equation (3.1). The characteristic equation is an algebraic equation of the nth. degree in q. The roots of this equation determine the possible values of q that allow (3.8) to be a solution of (3.6).

If the roots of (3.11) are the *distinct* numbers $q_1, \cdots q_n$, then the *general solution* of (3.6), may be written in the form,

$$(3.12) \qquad y_c(x) = A_1 q_1^x + A_2 q_2^x + \cdots + A_n q_n^x$$

where $A_1, A_2, \cdots A_n$ are *arbitrary constants*.

This solution is valid provided the roots q_i are *distinct*, regardless of whether they are real or complex.

If the characteristic equation (3.11) has a *multiple root* q_1 of the rth. order, it can be shown [7] that the part of the complementary function y_c that corresponds to this repeated root is,

$$(3.13) \qquad y_{c1}(x) = (C_1 + C_2 x + C_3 x^2 + \cdots + C_r x^{r-1}) q_1^x$$

where the C_i's are *arbitrary constants*.

The Particular Integral, $y_p(x)$.

The differential equations most frequently encountered in practice are linear and homogeneous with constant coefficients. The various methods for obtaining the particular integrals of linear equations with constant coefficients have their counterpart in the theory of difference equations. [5, 6, 7, 8].

The method to be used in determining the particular integral y_p depends on the nature of the function $V(x)$. To illustrate the general procedure, let the function $V(x)$ be of the form,

$$(3.14) \qquad V(x) = Ke^{mx}$$

The particular integral y_p in this case satisfies,

$$(3.15) \qquad P(E)y_p = Ke^{mx}$$

In this case, to obtain the particular integral assume a solution of the form,

$$(3.16) \qquad y_p = Be^{mx}$$

where B is to be determined. It is noticed that,

$$(3.17) \qquad E^k Be^{mx} = Ae^{mk}e^{mx}$$

It is therefore evident that,

$$P(E)Be^{mx} = Ae^{mx}P(e^m) \tag{3.18}$$

Hence if (3.16) is substituted into (3.15) the result is,

$$Be^{mx}P(e^m) = Ke^{mx} \tag{3.19}$$

Therefore if $P(e^m) \neq 0$, B is given by,

$$B = K/P(e^m), \tag{3.20}$$

and the *particular integral* of (3.15) is

$$y_p = Ke^{mx}/P(e^m) \tag{3.21}$$

In general, the particular integral of (3.1) may be expressed symbolically in the from,

$$y_p = \frac{1}{P(E)}V(x) = P^{-1}(E)V(x) \tag{3.22}$$

If $V(x)$ is a rational and integral function of x it is possible to expand $P^{-1}(E)$ in a series of ascending powers of Δ by the substitution $E = (1+\Delta)$ and operate with this series on $V(x)$. If $V(x)$ is a rational and integral function of x, the result of this operation will give a finite number of terms.

In some cases it is convenient to expand the operator $P^{-1}(E)$ into *partial fractions*. If this is done, the required particular integral may be obtained in terms of indefinite sums. [7].

IV. *Applications of Difference Equations to the Analysis of a Uniform Electric Network.*

An interesting and typical technical problem that illustrates the use of difference equations, is to determine the current and potential distribution of the electric circuit shown in Figure 1, in the periodic steady alternating state.

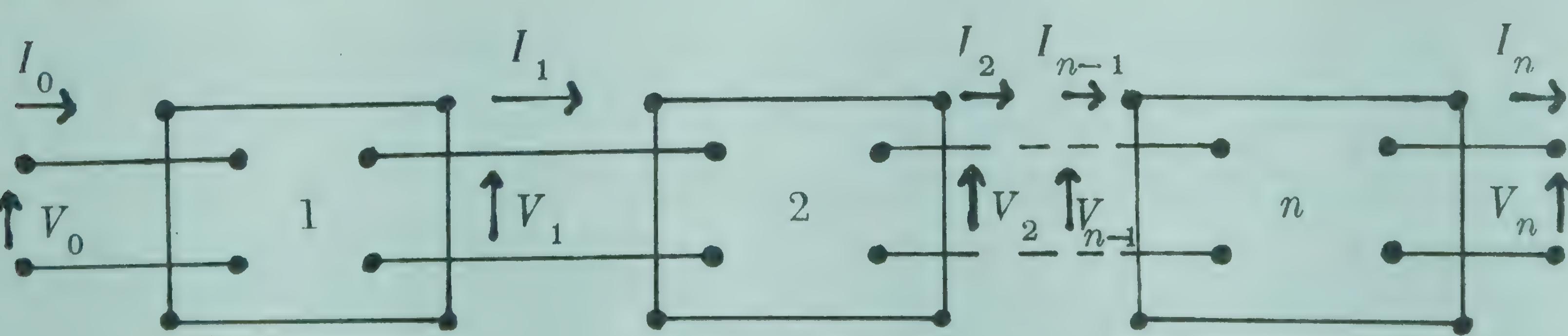
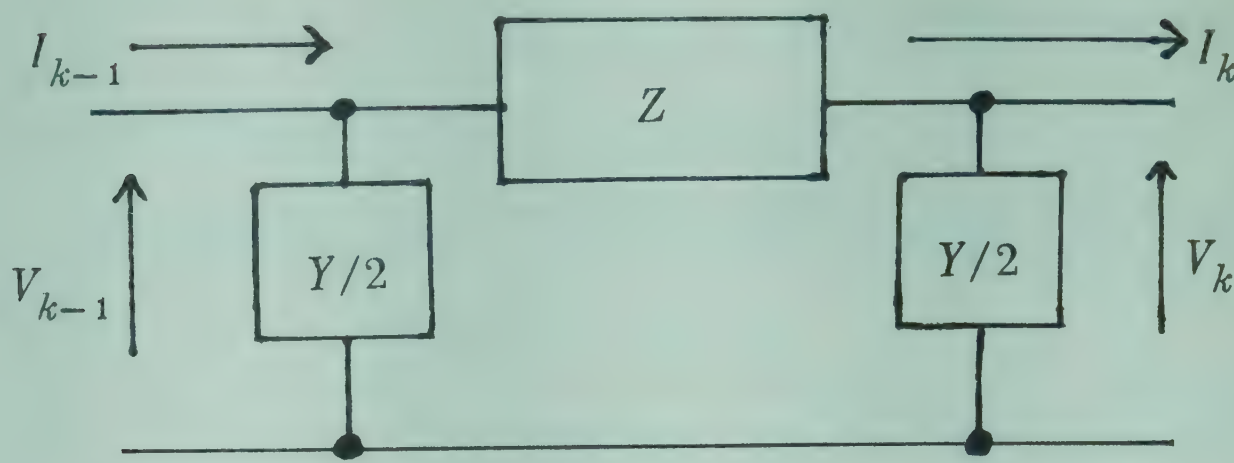


Figure 1.

The circuit of Figure 1 is a cascade connection of n boxes. Each box contains the *Pi* Network of impedances Z and admittances $Y/2$, shown in Figure 2.



Circuit Inside the kth. Box
Figure 2.

Each box of the network of Figure 1 contains the *Pi* Network of Figure 2. This consists of a symmetrical arrangement of two shunt admittances $Y/2$ and a series impedance Z.

If the sending end potential V_0 and the sending end current I_0 are specified, it is required to determine the potential and current distribution of the uniform network. If Kirchhoff's first and second laws are applied to the network of Figure 2, the following simultaneous difference equations are obtained:

(4.1) $$I_{k-1} = I_k + \frac{1}{2} Y(V_k + V_{k-1})$$

(4.2) $$V_{k-1} = V_k + Z(I_{k-1} - \frac{Y}{2} V_{k-1})$$

In order to solve the simultaneous system of difference equations (4.1) and (4.2), it is necessary to separate the variables. If (4.2) is solved for I_{k-1}, the result is,

(4.3) $$I_{k-1} = \frac{1}{Z}(V_{k-1} - V_k) + \frac{Y}{2} V_{k-1}$$

If this equation is used to eliminate I_{k-1} and I_k in (4.1), the result is,

(4.4) $$V_{k+1} - (2+ZY)V_k + V_{k-1} = 0$$

This is a homogeneous difference equation of the second order of the type discussed in Section III. Its solution is of the form,

(4.5) $$V_k = Aq^k$$

where A is an arbitrary constant and q satisfies the characteristic equation,

(4.6) $$q^2 - (2+ZY)q + 1 = 0$$

The two roots of this equation, satisfy the equations,

(4.7) $$q_1 - q_2 = 1$$

(4.8) $$q_1 + q_2 = (2 + ZY)$$

Let

(4.9) $$q_1 = e^a, \quad q_2 = e^{-a},$$

then

(4.10) $$q_1 + q_2 = (2+ZY) = 2\cosh(a)$$

The Propagation Constant of the Network.

(4.11) $$a = \cosh^{-1}(1+ZY/2) = \text{Propagation Constant.}$$

(4.12) $$V_k = A_1 e^{ak} + A_2 e^{-ak}$$

(4.13) $$I_k = \frac{T}{Z}(V_k - V_{k+1}) + \frac{Y}{2}V_k = \frac{1}{Z_0}[-A_1 e^{ak} + A_2 e^{-ak}]$$

The Characteristic Impedance of the Network. $= Z_0$

(4.14) $$Z_0 = \frac{2}{Y}\tanh(a/2)$$

Determination of the Arbitrary Constants.

The two arbitrary constants A_1 and A_2 may be obtained from the boundary conditions.

(4.15) $$V_0 = A_1 + A_2$$

(4.16) $$I_0 = \frac{1}{Z_0}(-A_1 + A_2)$$

Hence,

(4.17) $$A_1 = \frac{1}{2}(V_0 - Z_0 I_0)$$

(4.18) $$A_2 = \frac{1}{2}(V_0 + Z_0 I_0)$$

If these values of A_1 and A_2 are substituted into (4.12) and (4.13) the result may be expressed in the form,

(4.19) $$V_k = V_0 \cosh(ka) - Z_0 I_0 \sinh(ka)$$

(4.20) $$I_k = -V_0 \frac{\sinh(ka)}{Z_0} + I_0 \cosh(ka)$$

$$, \quad k = 0, 1, 2, \dots n$$

If

$$(4.21)\qquad [T] = \begin{bmatrix} \cosh(a) & Z_0\sinh(a) \\ \dfrac{\sinh(a)}{Z_0} & \cosh(a) \end{bmatrix} = \text{Transmission Matrix.}$$

$$(4.21)\qquad (y)_k = \begin{bmatrix} V_k \\ I_k \end{bmatrix}$$

Then,

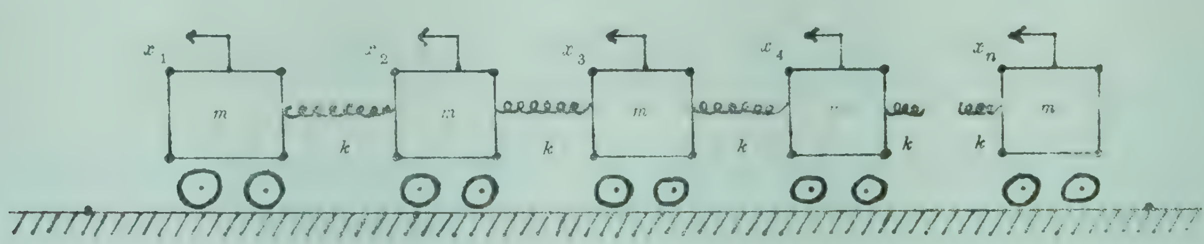
$$(4.22)\qquad (y)_k = [T]^{-k}(y)_0$$

and

$$(4.23)\qquad (y)_0 = [T]^n(y)_n, \qquad (y)_n = [T]^{-n}(y)_0$$

V. *Longitudinal Oscillations of a Long Uniform Train.*

Another typical example of the use of difference equations is the application to the dynamical problem of determining the resonant frequencies and modes of oscillations of a chain of n cars coupled by springs that satisfy hooke's law as shown in the Figure 3.



Idealized Uniform Train of n Cars

Figure 3.

Figure 3 represents an idealized train of n cars of equal mass m coupled to each other by identical springs of spring constant k. The effects of friction are neglected. The simplest method of analyzing the behavior of this system is by the electrical analog method which makes use of the results of Section IV. [See 10 in the Bibliography]

In this section, the dynamical system of Figure 3 will be analyzed by the use of a chain of differential equations and difference equations.

If Newton's second law of motion is applied to the rth. car in the chain of cars of Figure 3, the following differential equation is obtained.

$$(5.1)\qquad mx_r'' + k(x_r - x_{r-1}) + k(x_r - x_{r+1}) = 0, \qquad r = 2, 3, \ldots (n-1)$$

In order to study the oscillations of the system, assume a solution of (5.1) of the form,

$$(5.2)\qquad x_r = y_r \sin(\omega t + \theta)$$

where y_r is an amplitude function and ω is an angular frequency to be determined. θ is an arbitrary constant.

If (5.2) is substituted into (5.1) and the result divided by $\sin(\omega t+\theta)$, one obtains.

$$y_{r+1}-2\theta y_r+y_{r-1}=0 \tag{5.3}$$

where

$$\theta=(1-\omega^2 m/2k) \tag{5.4}$$

To solve the difference equation (5.3) assume,

$$y_r = Aq^r \tag{5.5}$$

If this solution is substituted into (5.3) the following *characteristic equation* is obtained:

$$q^2-2\theta q+1=0 \tag{5.6}$$

Let the two roots of this equation be q_1 and q_2 and let,

$$q_1=e^a, \qquad q_2=e^{-a} \tag{5.7}$$

Therefore,

$$(q_1+q_2)/2=(e^a+e^{-a})/2=\cosh(a)=\theta \tag{5.8}$$

The solution of the difference equation (5.3) may be written in the convenient form,

$$y_r=B_1\cosh(ar)+B_2\sinh(ar), \quad r=2,3,4,\cdots(n-1) \tag{5.9}$$

where B_1 and B_2 are arbitrary constants.

The Boundary Conditions.

In order to obtain the constants B_1 and B_2, use must be made of the boundary conditions at the front and rear of the train. It is realized that there are no cars corresponding to $r=0$ and to $r=(n+1)$. In order to extend the solution (5.9) to take into account $r=0$ and $r=(n+1)$ assume that there are phantom cars that correspond to these indices. Since the front and rear of the train is able to move freelù, there cannot be any strain in the phantom springs connecting the car $r=0$ to $r=1$ or the spring that is supposed to connect the car $r=n$ to the car $r=(n+1)$. This condition may be expressed mathematically by requiring that,

$$y_0=y_1 \tag{5.10}$$

$$y_n=y_{n+1} \tag{5.11}$$

and (5.9) can be used as if it were valid for $r=0$ and $r=(n+1)$.

The boundary conditions (5.10) and (5.11) lead to the two equations.

(5.12) $$B_1 = B_1\cosh(a) + B_2\sinh(a)$$

(5.13) $$B_1\cosh(an) + B_2\sinh(an) = B_1\cosh[a(n+1)] + B_2\sinh[a(n+1)]$$

B_2 can be expressed in terms of B_1 by (5.12) in the form,

(5.14) $$B_2 = -\tanh(a/2)B_1$$

If this expression for B_2 is substituted into (5.13) and some algebraic reductions are made, the result is,

(5.15) $$B_1\tanh(a/2)\sinh(an) = 0$$

For a non-trivial solution, we must have $B_1 = 0$. Hence the possible values of a must satisfy the transcendental equation,

(5.16) $$\tanh(a/2)\sinh(an) = 0$$

This equation is satisfied provided that,

(5.17) $$\sinh(an) = 0,$$

or

(5.18) $$a = a_s = j\pi s/n, \qquad s = 0, 1, 2, 3, \cdots (n-1). \quad j = (-1)^{1/2}$$

If s is given the value $s = n$, (5.16) leads to an indeterminate form and it is not satisfied.

The Resonant or Natural Frequencies of the Train.

To determine the possible natural angular frequencies of the system, we may use (5.8) written in the form,

(5.19) $$\cosh(a) = \cosh(a_s) = \theta = (1-\omega_s^2 m/2k) = \cosh(j\pi s/n)$$

Hence,

(5.20) $$\omega_s^2 m/2k = 1-\cosh(j\pi s/n) = 1-\cos(\pi s/n) = 2\sin^2(\pi s/2n)$$

Therefore,

(5.21) $$\omega_s = 2(k/m)^{1/2}\sin(\pi s/2n), \qquad s = 0, 1, 2, 3, \cdots (n-1)$$

These are the n natural angular frequencies of the principal modes of oscillation of the train. The frequency $\omega_0 = 0$, corresponds to a rigid body translation of the entire train.

The Natural Modes of Oscillation of the Train.

To determine the natural modes of oscillation of the system, write (5.9) in the form,

(5.22) $$y_r = B_1\cosh(ar) + B_2\sinh(ar)$$
$$= B_1[\cosh(ar)-\tanh(a/2)\sinh(ar)], \quad r = 1, 2, 3, \cdots n$$

To each value of $a = a_s = j\pi s/n$, there corresponds an amplitude function,

$$y_{rs} = B_s \phi(r, s), \tag{5.23}$$

where

$$\phi(r, s) = [\cos(\pi rs/n) - \tan(\pi s/2n)\sin(\pi rs/n)] \tag{5.24}$$
$$r = 1, 2, 3, \cdots n \quad \text{(cars)}$$
$$s = 0, 1, 2, \cdots (n-1) \quad \text{(modes)}$$

The general solution of the chain of differential equations (5.1) with the appropriate boundary conditions can now be obtained by substituting (5.23) into (5.2) and summing over all the possible modes s.

The general solution of the chain of equations (5.1), is therefore,

$$x_r = \sum_{s=0}^{s=(n-1)} B_r[\cos\frac{(\pi rs)}{n} - \tan\frac{(\pi s)}{2n}\sin\frac{(\pi rs)}{n}]\sin(\omega_s t + \theta_s) \tag{5.25}$$

The index r refers to the car number and the index s pertains to the mode. The quantities B_r and θ_s are the $2n$ arbitrary constants of the general solution. They may be obtained if the initial displacements $x_r(0)$ and the initial velocities $x_r'(0)$ of each car is given.

VI. *Difference Equation Method for Solving Partial Differential and Ordinary Differential Equations.*

There exists a vast literature in which the solution of ordinary and differential equations may be solved by the device of transforming the *differential equation* into an approximate *difference equation*. [1, 2, 3, 11]

From the definition of the derivative operator D and the operators Δ, ∇, and δ, it is easily seen that the following operators are approximations for the derivative operator D, provided h is sufficiently small. As $h \to 0$ the right members of these equations tend to D.

$$D \doteq \frac{\Delta}{h} \tag{6.1}$$

$$D \doteq \frac{\nabla}{h} \tag{6.2}$$

$$D \doteq \frac{\delta}{h} \tag{6.3}$$

To illustrate the use of these approximations to the operator D, let it be required to solve the differential equation,

$$\frac{d^2y}{dx^2} = x, \qquad \text{or} \qquad D^2 y = x. \tag{6.4}$$

Let the *initial conditions* be

$$y(0) = 0, \quad \text{and} \quad Dy(0) = 1 \tag{6.5}$$

For a sufficiently small h, we have approximately,

$$D^2y = \frac{\delta^2 y}{h^2} = \frac{(E^{1/2}-E^{-1/2})^2 y}{h^2} = \frac{(E-2+E^{-1})y(x)}{h^2} \tag{6.6}$$

Hence the differential equation (6.4) may be replaced by the approximate difference equation,

$$\frac{y(x+h)-2y(x)+y(x-h)}{h^2} = x \tag{6.7}$$

The initial conditions can be replaced by,

$$y(0) = 0, \quad \frac{y(h)-y(0)}{h} = 1 \tag{6.8}$$

Let the difference equation be required to be satisfied at the successive points, $x_1 = h,\ x_2 = 2h,\ \cdots\ x_k = kh$. Then the difference equation (6.7) may be written in the form,

$$y_{k+1} - 2y_k + y_{k-1} = kh^3, \quad (k = 1, 2, \cdots) \tag{6.9}$$

where

$$y_k = y(x_k) = y(kh). \tag{6.10}$$

The initial conditions take the form,

$$y_0 = 0, \quad y_1 - y_0 = h \tag{6.11}$$

The solution of (6.9) is

$$y_k = (h-h^3/6)k + h^3k^3/6 \tag{6.12}$$

This corresponds to,

$$y(x_k) = x_k + \frac{1}{6}x_k^3 - \frac{h^2}{6}x_k \tag{6.13}$$

The exact solution is,

$$y(x) = x + x^3/6 \tag{6.14}$$

If h is so small that h^2 can be neglected, (6.13) tends to (6.14).

Partial Differential Equations.

The Poisson differential equation will be taken as an example of the use of difference equations to solve partial differential equations. [3]. This expation has the form,

$$(D_x^2+D_y^2)u(x, y) = f(x, y) \tag{6.15}$$

In this case we have approximately,

(6.16) $$D_x^2 = (E_x - 2 + E_x^{-1})/h^2$$

(6.17) $$D_y^2 = (E_y - 2 + E_y^{-1})/h^2$$

where,

(6.18) $$E_x u(x, y) = u(x+h, y), \qquad E_y u(x, y) = u(x, y+h)$$

If the operators (6.16) and (6.17) are used, the partial differential equation (6.15) may be written in the approximate form,

(6.19) $$u(x+h,y) - 2u(x,y) + u(x-h,y) + u(x,y-h) - 2u(x,y) + u(x,y-h) = h^2 f(x,y)$$

The solution of this equation with prescribed boundary conditions is solved by Rosenbloom [3].

VII. *Additional Applications.*

Additional applications of Difference equations have been made to the study of intermittent control systems [12] and the analysis of systems with hereditary characteristics [13], as well as in certain problems in the theory of probability [14].

BIBLIOGRAPHY

1. Fox, L. "The Numerical Solution of Two-Point Boundary Problems in Ordinary Differential Equations", Oxford University Press, London, 1957.

2. Richtmyer, R. D., "Difference Methods for Initial Value Problems", Interscience Publishers, Inc., New York, 1957.

3. Beckenbach, E. F., "Construction and Applications of Conformal Maps", National Bureau of Standards, Applied Mathematics Series, 18, U. S. Government Printing Office, Washington, D. C., 1952.

4. Courant, R., Friedrichs, K. O. and Lewy, H., "Über die partiellen Differenzengleichungen der mathematischen Physik", Math. Annalen, vol. 100, p. 32, (1928).

5. Pipes, L. A., "Applied Mathematics for Engineers and Physicists", *Second Edition*, McGraw-Hill Book Company, Inc., New York, 1958.

6. Steffensen, J. F., "Interpolation", London, 1927.

7. Jordan, C., "Calculus of Finite Differences", Chelsea Publishing Co., New York, 1947.

8. Boole, G., "Treatise on the Calculus of Finite Differences", G. E. Stechert and Company, New York, 1946.

9. Bickley, W. J., "Difference and Associated Operators with Some Applications", Journal of Math. and Physics, Vol. 27, 1948, pp. 183-192.

10. Pipes, L. A., "Analysis of the Longitudinal Motions of Trains by the Electrical Analog", Jour. of Appl. Phys., Vol. 13, No. 12, Dec. 1942, pp. 780-786.

11. Hildebrand, F. B., "Methods of Applied Mathematics", Prentice-Hall, Inc., New York, 1952, Chapter II.

12. Tustin, A., "Automatic and Manual Control", Academic Press, New York,

1952.

13. Davis, H. T., "The Theory of Linear Operators", The Principia Press, Bloomington, Indiana, 1936.

14. Carslaw, H. S. and Jaeger, J. C., "Operational Methods in Applied Mathematics", Sec. Ed. Oxford Univ. Press, London, 1947.

University of California at Los Angeles

Editor, MATHEMATICS MAGAZINE, –

Success to Professor D. H. Hyers, and praise for his establishment of the S.I.M.N.T.! (See Mathematics Magazine, March-April, 1959.) Please enroll me as a member, and for my Habilitātationvorschlag I propose to replace "denumerable" and "enumerable" (which sounds too much like "innumerable") by "countable".

Let the Executive Committee also take under consideration words whose primary meanings do not suggest their mathematical significance, such as "field", "ring", "loop", "histogram", "regression", etc.

And can we get rid of the word "imaginary", so misleading to-day, and substitute the single letter, "i" (or "j"), both as noun and as adjective, with the definition $i^2 = -1$, and drop the $\sqrt{-1}$?

William R. Ransom
Tufts University

EULERIAN NUMBERS AND POLYNOMIALS

L. Carlitz

1. Introduction. Following Euler [6, pp. 487-491], we may put

$$\frac{1-\lambda}{e^x-\lambda} = \sum_{n=0}^{\infty} H_n \frac{x^n}{n!} \qquad (\lambda \neq 1), \tag{1.1}$$

where $H_n = H_n[\lambda]$ is a rational function of λ; indeed

$$R_n = R_n[\lambda] = (\lambda-1)^n H_n[\lambda] \tag{1.2}$$

is a polynomial in λ of degree $n-1$ with integral coefficients. If we put

$$R_n = \sum_{s=1}^{n} A_{ns}\lambda^{s-1} \qquad (n \geqq 1), \tag{1.3}$$

then the first few values of A_{ns} are given by the following table, where n denotes the row and s the column;

(1.4)

1					
1	1				
1	4	1			
1	11	11	1		
1	26	66	26	1	
1	57	302	302	57	1

Alternatively, Worpitzky [15] showed that the A_{ns} may be defined by means of

$$x^n = \sum_{s=1}^{n} A_{ns} \binom{x+s-1}{n}. \tag{1.5}$$

The rational functions H_n were studied in great detail by Frobenius [7], who was particularly interested in their relationship to the Bernoulli numbers. More recently Vandiver [14] has also made use of this relationship to obtain new properties of the Bernoulli numbers. Other recent occurrences are [1], [2], [12]; generalizations occur in [4], [5], [13]. In view

of the long history of H_n and A_{ns} it is rather curious that, on the whole, these quantities are not very well known. Indeed an examination of Mathematical Reviews for the past ten years will indicate that they have been frequently rediscovered. Actually there is no detailed discussion of H_n in any book. On the other hand, Riordan, in his recent book [11], does develop a few basic properties and indicates the connection of A_{ns} with certain combinational problems.

The present paper is mainly expository. We include numerous properties of H_n and the related polynomial

$$H_n(u \mid \lambda) = \sum_{r=0}^{n} \binom{n}{r} H_r u^{n-r}, \tag{1.6}$$

indicate the connection with Bernoulli numbers and polynomials and finally obtain some arithmetic properties of H_n. For the combinatorial applications the reader is referred to Riordan's book [11]. The H_n also occur in certain criteria for Fermat's last theorem; this is discussed at length in Bachmann's book [3] and will not be considered in the present paper.

2. The defining relation (1.1) is evidently equivalent to

$$(H+1)^n = \lambda H_n \qquad (n>0), \quad H_0 = 1, \tag{2.1}$$

where after expansion of the left member, H^n is replaced by H_n; we shall use this convention frequently. If $f(z)$ is an arbitrary polynomial in z, (2.1) implies

$$f(H+1) = \lambda f(H) + (1-\lambda) f(0). \tag{2.2}$$

In particular, for $f(z) = \binom{z}{m}$, we get

$$\binom{H+1}{m} = \lambda \binom{H}{m} \qquad (m \geqq 1), \tag{2.3}$$

which implies

$$(\lambda-1)\binom{H}{m} = \binom{H}{m-1} \qquad (m \geqq 1). \tag{2.4}$$

Repeated application of (2.3) gives

$$(\lambda-1)^r \binom{H}{m} = \binom{H}{m-r} \qquad (m \geqq r);$$

in particular we have

$$(\lambda-1)^m \binom{H}{m} = 1. \tag{2.5}$$

It also follows from (2.2) and (2.3) that

$$\binom{H+r}{m} = \lambda^r \binom{H}{m} \qquad (m \geqq r); \tag{2.6}$$

in particular by (2.5)

$$\binom{H+m}{m} = \frac{\lambda^m}{(1-\lambda)^m}. \tag{2.7}$$

Again, if $f(z)$ is an arbitrary polynomial of degree n, we recall that

$$f(z) = \sum_{r=0}^{n} \binom{z}{r} \Delta^r f(0),$$

where

$$\Delta f(z) = f(z+1) - f(z), \quad \Delta^r f(z) = \Delta^{r-1} f(z+1) - \Delta^{r-1} f(z).$$

Using (2.5) we get

$$f(H) = \sum_{r=0}^{n} (\lambda-1)^{-r} \Delta^r f(0). \tag{2.8}$$

Since

$$\Delta^r f(0) = \sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} f(s),$$

we may write

$$f(H) = \sum_{r=0}^{n} (\lambda-1)^{-r} \sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} f(s). \tag{2.9}$$

It is perhaps of interest to mention that (2.9) can also be obtained as follows from (1.1). For $|x|$ sufficiently small we have

$$\frac{1-\lambda}{e^x-\lambda} = \left(1 - \frac{e^x-1}{\lambda-1}\right)^{-1} = \sum_{r=0}^{\infty} (\lambda-1)^{-r} (e^x-1)^r$$

$$= \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{r=0}^{n} (\lambda-1)^{-r} \sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} s^n .$$

so that

$$(2.10)\qquad H_n = \sum_{r=0}^{n} (\lambda-1)^{-r} \sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} s^n$$

and (2.9) follows at once.

Returning again to (1.1) it is easily verified that

$$(2.11)\qquad R_n[\lambda] = \lambda^{n-1} R_n[\lambda^{-1}];$$

also differentiation yields

$$(2.12)\qquad R_{n+1} = (n+1)\lambda R_n + (1-\lambda)\frac{d}{d\lambda}(\lambda R_n),$$

where R_n is defined by (1.2). From (2.11) and (2.12) we immediately obtain

$$(2.13)\qquad A_{ns} = A_{n,n-s+1}$$

and

$$(2.14)\qquad A_{n+1,s} = sA_{n,s-1} + (n-s+2)A_{ns}.$$

By means of (2.14) one can easily extend the table (1.4). A convenient check is furnished by the formula

$$R_n[1] = \sum_{s=1}^{n} A_{ns} = n! \qquad (n \geq 1).$$

We also note that $R_{2n}[-1] = 0$ for $n \geq 1$.

Frobenius remarks that it follows from (2.12) that the $n-1$ roots of $R_n[\lambda] = 0$ are real, negative and distinct; also for each root λ_0 the reciprocal λ_0 is also a root. Moreover the roots of $R_{n+1}[\lambda] = 0$ are separated by the roots of $R_n[\lambda] = 0$. In the next place, by (2.10)

$$R_n = \sum_{r=0}^{n} (\lambda-1)^{n-r} \sum_{j=0}^{r} (-1)^{r-j} \binom{r}{j} j^n$$

$$= \sum_{r=0}^{n} \sum_{s=0}^{n-r} (-1)^{n-r-s} \binom{n-r}{s} \lambda^s \sum_{j=0}^{r} (-1)^{r-j} \binom{r}{j} j^n$$

$$= \sum_{s=0}^{n} \lambda^s \sum_{j=0}^{n-s} (-1)^{n-s-j} j^n \sum_{r=j}^{n-s} \binom{n-r}{s}\binom{r}{j}$$

$$= \sum_{s=0}^{n} \lambda^s \sum_{j=0}^{n-s} (-1)^{n-s-j} \binom{n+1}{s+j+1} j^n.$$

Hence, by (1.3) and (2.13), we get

$$A_{ns} = \sum_{j=0}^{s} (-1)^j \binom{n+1}{j} (s-j)^n, \tag{2.15}$$

a formula due to Euler.

Since

$$x^n = \sum_{r=0}^{n} \binom{x}{r} \Delta^r f(0) = \sum_{r=0}^{n} \Delta^{n-r} \binom{x}{n} \Delta^r f(0),$$

it follows from (2.10) and (2.15)

$$x^n = R_n(1+\Delta)\binom{x}{n} = \sum_{s=1}^{n} A_{ns}(1+\Delta)^{s-1}\binom{x}{n} = \sum_{s=1}^{n} A_{ns}\binom{x+s-1}{n},$$

which establishes (1.5). This proof is taken from Frobenius.

3. We now consider the polynomial

$$H_n(u) = H_n(u \mid \lambda) = (u+H)^n \tag{3.1}$$

defined by (1.6). We evidently have the generating function

$$\frac{1-\lambda}{e^x-\lambda} e^{xu} = \sum_{n=0}^{\infty} \frac{x^n}{n!} H_n(u). \tag{3.2}$$

It follows from (3.2) that

$$H_n(u+1) - \lambda H_n(u) = (1-\lambda)u^n \tag{3.3}$$

Moreover (3.3) uniquely determines the polynomial $H_n(u)$. If $f(z)$ is an arbitrary polynomial in z then by (3.3)

$$f(u+1+H) - \lambda f(u+H) = (1-\lambda) f(u),$$

from which it follows that for given $f(z)$ the difference equation

$$g(u+1) - \lambda g(u) = (1-\lambda)f(u) \tag{3.4}$$

has the unique solution

$$g(u) = g(u \mid \lambda) = f(u+H).$$

It follows at once from (3.2) that

$$H_n'(u) = nH_{n-1}(u) \tag{3.5}$$

and generally

$$H_n^{(r)}(u) = \frac{n!}{(n-r)!}H_{n-r}(u) \qquad (n \geqq r),$$

which implies

$$H_n(u+v) = \sum_{r=0}^{n} \binom{n}{r} u^{n-r} H_n(v). \tag{3.6}$$

If we differentiate (3.2) with respect to λ we get

$$H_{n+1}(u \mid \lambda) + \lambda\frac{\partial}{\partial\lambda}H_n(u \mid \lambda) = \left(u - \frac{1}{1-\lambda}\right)H_n(u \mid \lambda), \tag{3.7}$$

which reduces to (2.12) for $u = 0$. We also note that (3.2) implies

$$H_n(1-u \mid \lambda^{-1}) = (-1)^n H_n(u, \lambda). \tag{3.8}$$

We remark that by (2.10) and (3.1) we have

$$H_n(u) = \sum_{r=0}^{n} (\lambda-1)^{-r} \sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} (u+s)^n. \tag{3.9}$$

Again, if we put

$$f_n(u, m) = \sum_{j=0}^{m-1} (u+j)^n \lambda^{m-1-j} \tag{3.10}$$

then it follows from (3.3) that

$$f_n(u, m) = \frac{H_n(u+m) - \lambda^m H_n(u)}{1-\lambda}. \tag{3.11}$$

The polynomial $f_n(0, m)$ is usually called a Mirimonoff polynomial; a more

general polynomial is discussed by Vandiver [14], see also Bachmann [1, p. 117].

To get a *multiplication* theorem for $H_n(u)$ we consider

$$\sum_{n=0}^{\infty}\frac{x^n}{n!}\sum_{r=0}^{m-1}\lambda^{m-1-r}H_n\left(u+\frac{r}{m}\mid\lambda^m\right) = \frac{1-\lambda^m}{e^x-\lambda^m}e^{xu}\sum_{r=0}^{m-1}\lambda^{m-1-r}e^{rx/m}$$

$$= \frac{1-\lambda^m}{e^{x/m}-\lambda}e^{xu}$$

$$= \frac{1-\lambda^m}{1-\lambda}\sum_{n=0}^{\infty}\frac{x^n m^{-n}}{n!}H_n(mu\mid\lambda),$$

which yields

$$m^n\sum_{r=0}^{m-1}\lambda^{m-1-r}H_n\left(u+\frac{r}{m}\mid\lambda^m\right) = \frac{1-\lambda^m}{1-\lambda}H_n(mu\mid\lambda). \tag{3.12}$$

An interesting special case of (3.12) is obtained by taking $\lambda = \zeta$, where $\zeta^{m-1} = 1$, $\zeta \neq 1$. Then (3.12) reduces to

$$m^n\sum_{r=0}^{m-1}\zeta^{-r}H_n\left(u+\frac{r}{m}\mid\zeta\right) = H_n(u\mid\zeta). \tag{3.13}$$

Nielsen [9, p. 54] has proved that the multiplication theorems for the Bernoulli and Euler polynomials characterize those polynomials. Suppose now that $f(u\mid\lambda)$ is a polynomial in u of degree n that satisfies the equation

$$m^n\sum_{r=0}^{m-1}\lambda^{m-1-r}f\left(u+\frac{r}{m}\mid\lambda^m\right) = \frac{1-\lambda^m}{1-\lambda}f(mu\mid\lambda) \tag{3.14}$$

for some value of $m>1$. Put

$$f(u\mid\lambda) = \sum_{s=0}^{n}A_s[\lambda]H_s(u\mid\lambda).$$

Then by (3.12) and (3.14)

$$\sum_{s=0}^{n} A_s[\lambda^m] m^{n-s} H_s(mu \mid \lambda) = \sum_{s=0}^{n} A_s[\lambda] H_s(mu \mid \lambda).$$

This requires that

$$m^{n-s} A_s[\lambda^m] = A_s[\lambda] \qquad (0 \leqq s \leqq n). \tag{3.15}$$

Now assume that (3.14) holds for *two* values of $m > 1$, say m_1 and m_2. Then it is clear that (3.15) becomes

$$A_n[\lambda^m] = A_n[\lambda], \quad A_s[\lambda] = 0 \qquad (0 \leqq s < n).$$

Therefore

$$f(u \mid \lambda) = A_n[\lambda] H_n(u \mid \lambda).$$

For (3.13) the situation is somewhat simpler. If $g(u \mid \zeta)$ is a polynomial in u satisfying the equation

$$m^n \sum_{r=0}^{m-1} \zeta^{-r} g\left(u + \frac{r}{m} \mid \zeta\right) = g(mu \mid \zeta), \tag{3.16}$$

where $\zeta^{m-1} = 1$, $\zeta \neq 1$, then if we put

$$g(u \mid \zeta) = \sum_{s=0}^{n} A_s(\zeta) H_n(u \mid \zeta)$$

and assume that (3.16) holds for *one* value of $m > 1$, then it follows readily that

$$g(u \mid \zeta) = A_n(\zeta) H_n(u \mid \zeta),$$

where $A_n(\zeta)$ is arbitrary.

It may be of interest to mention also an *addition* theorem satisfied by $H_n(u \mid \lambda)$. Since

$$\frac{1-\lambda}{e^x - \lambda} e^{xu} \frac{1+\lambda}{e^x + \lambda} e^{xv} = \frac{1-\lambda^2}{e^{2x} - \lambda^2} e^{x(u+v)},$$

it follows at once that

$$\sum_{r=0}^{n} \binom{n}{r} H_r(u \mid \lambda) H_{n-r}(v \mid -\lambda) = H_n(u+v \mid \lambda^2). \tag{3.17}$$

We note also that from the identity

$$\frac{(1-\lambda)(1+\lambda)}{e^x-\lambda}-\frac{(1-\lambda)(1+\lambda)}{e^x+\lambda}=\frac{2\lambda(1-\lambda^2)}{e^{2x}-\lambda^2}$$

follows

$$(1+\lambda)H_n(u\mid\lambda)-(1-\lambda)H_n(u\mid-\lambda)=2^{n+1}\lambda H_n\left(\frac{u}{2},\lambda^2\right), \tag{3.18}$$

while from

$$\frac{2\lambda-1}{(e^x-\lambda)(e^x-1+\lambda)}=\frac{1}{e^x-\lambda}-\frac{1}{e^x-1+\lambda}$$

follows

$$(2\lambda-1)\sum_{r=0}^{n}\binom{n}{r}H_r(u\mid\lambda)H_{n-r}(v\mid 1-\lambda)=\lambda H_n(u+v\mid\lambda)-(1-\lambda)H_n(u+v\mid 1-\lambda). \tag{3.19}$$

4. It is familiar that the Bernoulli polynomial $B_n(u)$ may be defined by [10, Chapter 2]

$$\frac{xe^{xu}}{e^x-1}=\sum_{n=0}^{\infty}B_n(u)\frac{x^n}{n!}\,. \tag{4.1}$$

then

$$\sum_{n=0}^{\infty}\frac{x^n}{n!}\sum_{s=0}^{m-1}\zeta^{-rs}B_n\left(u+\frac{s}{m}\right)=\frac{xe^{xu}}{e^x-1}\sum_{s=0}^{m-1}\zeta^{-rs}e^{sx/m}$$

$$=\frac{xe^{xu}}{\zeta^{-r}e^{x/m}-1}=\frac{\zeta^r xe^{xu}}{e^{x/m}-\zeta^r},$$

where ζ is a primitive m-th root of unity and $m\nmid r$. This evidently implies

$$m^{n-1}\sum_{s=0}^{m-1}\zeta^{-rs}B_n\left(u+\frac{s}{m}\right)=\frac{n\zeta^r}{1-\zeta^r}H_{n-1}(mu\mid\zeta^r). \tag{4.2}$$

For $m\mid r$, on the other hand, we have the multiplication theorem for $B_n(x)$:

$$m^{n-1}\sum_{s=0}^{m-1}B_n\left(u+\frac{s}{m}\right)=B_n(mu) \tag{4.3}$$

Multiplying both sides of (4.2) by ζ^{rt}, summing over r and using (4.3), we get

$$m^n B_n\left(u+\frac{t}{m}\right) = B_n(mu) + n\sum_{r=1}^{m-1} \zeta^{rt}\frac{H_{n-1}(mu \mid \zeta^r)}{\zeta^{-r}-1}, \tag{4.4}$$

where $0 \leqq t < m$.

We recall that the Bernoulli function $\overline{B}_n(u)$ is defined by

$$\overline{B}_n(u) = B_n(u) \quad (0 \leqq u < 1), \qquad \overline{B}_n(u+1) = \overline{B}_n(u).$$

Similarly we define

$$\overline{H}_n(u \mid \zeta) = H_n(u \mid \zeta) \quad (0 \leqq u < 1), \qquad \overline{H}_n(u+1 \mid \zeta) = \zeta\overline{H}_n(u \mid \zeta), \tag{4.5}$$

where ζ is some root of unity $\neq 1$. With these definitions of $\overline{H}_n(u \mid \zeta)$ and $\overline{B}_n(x)$ it is easily verified that the formulas (3.8), (3.12), (3.13), (4.2), (4.3), (4.4) hold for the barred functions; in particular in (4.4) the restriction $0 \leqq t < m$ is no longer necessary.

We remark that for m even and $\zeta^r = -1$, (4.2) reduces to the known formula

$$m^n \sum_{s=0}^{m-1} (-1)^s B_n\left(u+\frac{s}{m}\right) = -\frac{n}{2}E_{n-1}(mu),$$

where $E_{n-1}(u)$ is the Euler polynomial of degree $n-1$.

For $\zeta^r = \omega$, where $\omega^2 + \omega + 1 = 0$, $m = 3$, $u = 0$, (4.2) becomes

$$\frac{n\omega}{1-\omega} H_{n-1}[\omega] = 3^n\left\{B_n + \omega^{-1}B_n\left(\tfrac{1}{3}\right) + \omega B_n\left(\tfrac{2}{3}\right)\right\}.$$

For n even, it is known that [10, p. 22]

$$B_n\left(\tfrac{1}{3}\right) = B_n\left(\tfrac{2}{3}\right) = \tfrac{1}{2}(3^{1-n}-1)B_n,$$

from which it follows that

$$\frac{2n\omega}{1-\omega} H_{n-1}[\omega] = (3^n-1)B_n \qquad (n \text{ even}). \tag{4.6}$$

On the other hand for n odd > 1 we get

$$nH_{n-1}[\omega] = 3^n\omega B_n\left(\tfrac{1}{3}\right) \tag{4.7}$$

Again for $\zeta^r = i$, $m = 4$, we get

$$\frac{ni}{1-i} H_{n-1}[i] = 4^{n-1}\left\{B_n - iB_n(\tfrac{1}{4}) - B_n(\tfrac{1}{2}) - iB_n(\tfrac{3}{4})\right\}.$$

It follows that

$$n(i-1)H_{n-1}[i] = 2^n(2^n-1)B_n \qquad (n \text{ even}), \tag{4.8}$$

$$n(i+1)H_{n-1}[i] = -2^{2n}B_n(\tfrac{1}{4}) \qquad (n \text{ odd} > 1). \tag{4.9}$$

For $\lambda = -1$ we have [10, p. 28]

$$H_{n-1}[-1] = 2^{1-n}C_{n-1} = 2(1-2^n)\frac{B_n}{n}. \tag{4.10}$$

5. We now obtain some congruences satisfied by H_n. If in (2.8) we take $f(z) = z^n(1-z^w)^r$, we get

$$H^n(1-H^w)^k = \sum_r (\lambda-1)^{-r} \sum_{s=0}^{n} (-1)^{r-s}\binom{r}{s} s^n(1-s^w)^k.$$

Assume that p is a prime such that

$$(p-1)p^{e-1} \mid w; \tag{5.1}$$

then by Fermat's theorem

$$s^n(1-s^w)^k \equiv 0 \left(\operatorname{mod}\ (p^n, p^{kw})\right).$$

We thus obtain (Frobenius)

$$\sum_{s=0}^{k} (-1)^{k-s}\binom{k}{s} H_{n+sw} \equiv 0 \left(\operatorname{mod}\ (p^n, p^{kw})\right) \tag{5.2}$$

valid for $k \geqq 0$, $n \geqq 0$, provided w satisfies (5.1). This result is referred to as Kummer's congruence for H_n. In (5.2) λ may be an indeterminate or an algebraic number such that $(p, 1-\lambda) = (1)$. In particular λ may be an l-th root of unity, where $l \neq p^f$.

By means of (5.2) it is proved in [5] that the coefficient A_{ns} of (1.3) satisfies the congruence

$$A_{n+b,\, s} \equiv A_{n,\, s} \pmod{p^e}, \tag{5.3}$$

where $p^{j-1} < s \leqq p^j$, $n \geqq e$, $b = p^{j+s-1}(p-1)$.

Another interesting result, also due to Frobenius, is

$$H_w \equiv \frac{\lambda^{p-1}-1}{\lambda^p-1} \pmod{p^e}, \tag{5.4}$$

where again w satisfies (5.1). Indeed we can prove a slightly more general result by using (3.11) with $n = w$, $u = r$, $m \equiv 0 \pmod{p^e}$. We get

$$\frac{1-\lambda^m}{1-\lambda} H_w(r) \equiv \sum_{\substack{j=0 \\ p\nmid(r+j)}}^{m-1} \lambda^{m-1-j} \equiv \sum_{j=0}^{m-1} \lambda^{m-1-j} - \sum_{i=0}^{\frac{m}{p}-1} \lambda^{m-1-r_0-ip}$$

$$\equiv \frac{\lambda^m-1}{\lambda-1} - \lambda^{p-1-r_0} \frac{\lambda^m-1}{\lambda^p-1} \pmod{p^e},$$

where the integer r_0 is defined by

$$r \equiv -r_0 \pmod{p} \qquad (0 \leqq r_0 < p). \tag{5.5}$$

We have finally

$$H_w(r) \equiv 1 - \lambda^{p-1-r_0} \frac{\lambda-1}{\lambda^p-1} \pmod{p^e}, \tag{5.6}$$

which reduces to (5.4) for $r = 0$.

The interesting congruence

$$R_{p-2}[1-\lambda] \equiv R_{p-2}[\lambda] \pmod{p} \tag{5.7}$$

is due to Mirimanoff [8]. It can be proved rapidly as follows. From (3.11) we get

$$(1-\lambda)R_{p-2}[\lambda] \equiv \sum_{j=0}^{p-2} (j+1)^{-1}\lambda^j,$$

so that

$$\frac{d}{d\lambda}\{\lambda(1-\lambda)R_{p-2}[\lambda]\} \equiv \frac{\lambda-\lambda^p}{(1-\lambda)}. \tag{5.8}$$

Replacing λ by $1-\lambda$, we get also

$$\frac{d}{d\lambda}(\lambda(1-\lambda)R_{p-2}[1-\lambda]) \equiv \frac{\lambda-\lambda^p}{(1-\lambda)}.$$

Consequently

$$\lambda(1-\lambda)R_{p-2}[1-\lambda] \equiv \lambda(1-\lambda)R_{p-2}[\lambda] + C,$$

where C is independent of λ. Clearly $C \equiv 0$ and (5.7) follows at once.

Again by (3.11)

$$(1-\lambda)^2 R_{p-3}[\lambda] \equiv \sum_{j=0}^{p-2} (j+1)^{-2}\lambda^j ,$$

so that

$$\frac{d}{d\lambda}\{\lambda(1-\lambda)^2 R_{p-3}[\lambda]\} \equiv (1-\lambda)R_{p-2}[\lambda] . \tag{5.9}$$

Then by (5.7)

$$(\lambda^p-\lambda)R_{p-2}[\lambda] \equiv \lambda^p(1-\lambda)R_{p-2}[\lambda] - \lambda(1-\lambda)^p R_{p-2}[1-\lambda] ;$$

now using (5.8) and (5.9), we get

$$-\tfrac{1}{2}\lambda^2(1-\lambda)^2 R_{p-2}^2[\lambda] \equiv \lambda^{p+1}(1-\lambda)^2 R_{p-3}[\lambda] + \lambda^2(1-\lambda)^{p+1} R_{p-3}[1-\lambda] ,$$

so that

$$-\tfrac{1}{2}R_{p-2}^2[\lambda] \equiv \lambda^{p-1}R_{p-3}[\lambda] + (1-\lambda)^{p-1}R_{p-3}[1-\lambda] \pmod p , \tag{5.10}$$

where of course $p>2$. This result also is due to Mirimanoff.

We note that from (2.1) and

$$\binom{p-1}{r} \equiv (-1)^r \pmod p$$

it follows that

$$\sum_{r=0}^{p-1} (-1)^r H_r \equiv \lambda H_{p-1} \pmod p , \tag{5.11}$$

$$p\sum_{r=1}^{p-1} \frac{(-1)^{r-1}}{r} H_r \equiv (\lambda-1)H_p - 1 \pmod{p^2} . \tag{5.12}$$

Finally from (3.11) follows

$$\sum_{r=1}^{p-1} \left(\frac{r}{p}\right)\lambda^{r-1} \equiv (1-\lambda)^k R_k \pmod p , \tag{5.13}$$

where $p = 2k+1$ and (r/p) is the Legendre symbol.

REFERENCES

1. T. M. Apostol, On the Lerch zeta function, Pacific Journal of Mathematics vol. 1(1951), pp. 161-167.

2. T. M. Apostol, Addendum to 'On the Lerch zeta function,' Pacific Journal of Mathematics, vol. 2(1952), p. 10.

3. P. Bachmann, Das Fermat problem in seiner bisherigen Entwicklung, Berlin and Leipzig, 1919.

4. L. Carlitz, q-Bernoulli and Eulerian numbers, Transactions of the American Mathematical Society, vol. 76(1954), pp. 332-350.

5. L. Carlitz and J. Riordan, Congruences for Eulerian numbers, Duke Mathematical Journal, vol. 20(1953), pp. 339-343.

6. L. Euler, Institutiones calculi differentialis, Petrograd, 1755.

7. G. Frobenius, Über die Bernoullischen und die Eulerschen Polynome, Sitzungsberichte der Preussische Akademie der Wissenschaften (1910), pp. 809-847.

8. D. Mirimanoff, L'équation indéterminée $x^l + y^l + z^l = 0$ et le criterium de Kummer, Journal fur die reine und angewandte Mathematik, vol. 128(1905), pp. 45-68.

9. N. Nielsen, Traité élémentaire des nombres de Bernoulli, Paris, 1923.

10. N. E. Norlund, Vorlesungen über Differenzenrechnung, Berlin, 1924.

11. J. Riordan, An Introduction to Combinational Analysis, New York, 1958.

12. J. Riordan, Triangular permutation numbers, Proceedings of the American Mathematical Society, vol. 2(1951), pp. 429-432.

13. E. B. Shanks, Iterated sums of powers of binomial coefficients, American Mathematical Monthly, vol. 58(1951), pp. 404-407.

14. H. S. Vandiver, An arithmetical theory of the Bernoulli numbers, Transactions of the American Mathematical Society, vol. 51(1942), pp. 502-531.

15. J. Worpitzky, Studien über die Bernoullischen und Eulerschen Zahlen, Journal für die reine und angewandte Mathematik, vol. 94(1883), pp. 203-232.

Duke University
Durham, N. C.

PROPORTIONAL METRICS IN N VARIABLES

H. Randolph Pyle

Introduction.

Suppose that we are given a surface S with metric $ds^2 = Edx^2 + 2Fdxdy + Gdy^2$ and a surface S' with metric $ds'^2 = E'dx'^2 + 2F'dx'dy' + G'dy'^2$. The point (x', y') of S' is related to the point (x, y) of S by the transformation $x' = \phi(x, y)$, $y' = \psi(x, y)$. The necessary and sufficient condition for the conformal mapping of S on S' is $ds'^2 = \lambda^2 ds^2$. We have shown in a previous paper (1) that this is equivalent to a set of linear relations between the partial derivatives of ϕ and ψ that reduce to the Cauchy-Riemann equations for isometric parameters. It is the purpose of this paper to find similar relations when the metrics are expressed in terms of n parameters instead of two. This will be accomplished by the use of matrix methods.

Conditions for proportional metrics.

We shall express the first metric as $ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} dx_i dx_j$ and the second as $ds'^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} dy_i dy_j$, and impose the condition that $ds'^2 = \lambda^2 ds^2$. The y's are related to the x's by the transformation $y_i = \phi_i(x_1, \cdots x_n)$, $(i = 1, \cdots n)$. We shall make the usual assumptions concerning the existence and the continuity of the partial derivatives of the ϕ_i, and that the Jacobian does not vanish. Let the symbol ϕ_{ik} stand for $\dfrac{\partial \phi_i}{\partial x_k}$, so that $\phi_{ik} = \dfrac{\partial \phi_i}{\partial x_k} = \dfrac{\partial y_i}{\partial x_k}$. The Jacobian J is the determinant of the matrix

$$\Delta = \begin{pmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2n} \\ \cdots \\ \phi_{n1} & \phi_{n2} & \cdots & \phi_{nn} \end{pmatrix}, \qquad J = |\Delta|.$$

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

$(a_{ij} = a_{ji},\ b_{ij} = b_{ji})$ be the matrices of the coefficients of the differentials in ds^2 and ds'^2 respectively. Designate the determinants of these matrices by $|A|$, $|B|$, and assume that $|A| \neq 0$, $|B| \neq 0$. Let $dX = (dx_1, dx_2, \cdots dx_n)$, $dY = (dy_1, dy_2, \cdots dy_n)$ be row vectors and $(dX)^{\mathrm{T}}$, $(dY)^{\mathrm{T}}$ be the corresponding column vectors.

Using matrix multiplication, we have $ds^2 = (dX)A(dX)^{\mathrm{T}}$ and $ds'^2 = (dY)B(dY)^{\mathrm{T}}$. Since $dy_i = \sum_{k=1}^{n} \phi_{ik} dx_k$, $(dY)^{\mathrm{T}} = \Delta(dX)^{\mathrm{T}}$ and $dY = (dX)\Delta^{\mathrm{T}}$. Then $ds'^2 = (dX)\Delta^{\mathrm{T}}B\Delta(dX)^{\mathrm{T}}$. When $ds'^2 = \lambda^2 ds^2$, $\Delta^{\mathrm{T}}B\Delta = \lambda^2 A$.

Let $\overline{\Phi}_{ij}$ represent the cofactor of ϕ_{ij} in J, and

$$\overline{\Delta} = \begin{pmatrix} \overline{\Phi}_{11} & \overline{\Phi}_{12} & \cdots & \overline{\Phi}_{1n} \\ \overline{\Phi}_{21} & \overline{\Phi}_{22} & \cdots & \overline{\Phi}_{2n} \\ \cdots \\ \overline{\Phi}_{n1} & \overline{\Phi}_{n2} & \cdots & \overline{\Phi}_{nn} \end{pmatrix} \qquad \Delta^{\mathrm{T}} = \begin{pmatrix} \phi_{11} & \phi_{21} & \cdots & \phi_{n1} \\ \phi_{12} & \phi_{22} & \cdots & \phi_{n2} \\ \cdots \\ \phi_{1n} & \phi_{2n} & \cdots & \phi_{nn} \end{pmatrix}$$

Then

$$\overline{\Delta}\Delta^{\mathrm{T}} = \begin{pmatrix} J & 0 & \cdots & 0 \\ 0 & J & \cdots & 0 \\ \cdots \\ 0 & 0 & \cdots & J \end{pmatrix} = JI$$

where I is the identity matrix. Now let $L_{ij} = \sum_{k=1}^{n} b_{ik}\phi_{kj}$, $\overline{L}_{ij} = \sum_{k=1}^{n} a_{kj}\overline{\Phi}_{ik}$ and

$$L = B\Delta = \begin{pmatrix} L_{11} & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22} & \cdots & L_{2n} \\ \cdots \\ L_{n1} & L_{n2} & \cdots & L_{nn} \end{pmatrix}, \quad \overline{L} = \overline{\Delta}A = \begin{pmatrix} \overline{L}_{11} & \overline{L}_{12} & \cdots & \overline{L}_{1n} \\ \overline{L}_{21} & \overline{L}_{22} & \cdots & \overline{L}_{2n} \\ \cdots \\ \overline{L}_{n1} & \overline{L}_{n2} & \cdots & \overline{L}_{nn} \end{pmatrix}.$$

Since $\Delta^{\mathrm{T}}B\Delta = \lambda^2 A$, $\overline{\Delta}(\Delta^{\mathrm{T}}B\Delta) = \overline{\Delta}(\lambda^2 A)$, or $(\overline{\Delta}\Delta^{\mathrm{T}})(B\Delta) = \lambda^2(\overline{\Delta}A)$, and $JL = \lambda^2\overline{L}$. Also $|\Delta^{\mathrm{T}}||B||\Delta| = |\lambda^2 A|$. But $|\Delta^{\mathrm{T}}| = |\Delta| = J$, and $|\lambda^2 A| = \lambda^{2n}|A|$, so that $J^2|B| = \lambda^{2n}|A|$, and $\lambda^2 = \left(\frac{J^2|B|}{|A|}\right)^{1/n}$.

Since $JL = \lambda^2\overline{L}$, if we call $\mu = \frac{J}{\lambda^2} = \frac{|A|^{1/n} J^{(n-2)/n}}{|B|^{1/n}}$, we have $\overline{L} = \mu L$, and $\overline{L}_{ij} = \mu L_{ij}$, or

$$\sum_{k=1}^{n} a_{kj}\overline{\Phi}_{ik} = \mu \sum_{k=1}^{n} b_{ik}\phi_{kj}$$

as the necessary conditions for $ds'^2 = \lambda^2 ds^2$.

Since each of these steps is reversible when J, $|A|$, $|B| \neq 0$, these conditions are also sufficient.

Special cases.

1. When $n = 2$.

$$ds^2 = E dx_1^2 + 2F dx_1 dx_2 + G dx_2^2$$

$$ds'^2 = E' dy_1^2 + 2F' dy_1 dy_2 + G' dy_2^2$$

$$A = \begin{pmatrix} E & F \\ F & G \end{pmatrix} \quad B = \begin{pmatrix} E' & F' \\ F' & G' \end{pmatrix} \quad |A| = H^2, \quad |B| = H'^2$$

$$y_1 = \phi(x_1, x_2), \quad y_2 = \psi(x_1, x_2), \quad \Delta = \begin{pmatrix} \phi_x & \phi_y \\ \psi_x & \psi_y \end{pmatrix}, \quad \overline{\Delta} = \begin{pmatrix} \psi_y & -\psi_x \\ -\phi_y & \phi_x \end{pmatrix}$$

$$L_{11} = E'\phi_x + F'\psi_x, \quad L_{12} = E'\phi_y + F'\psi_y, \quad \overline{L}_{11} = E\psi_y - F\psi_x, \quad \overline{L}_{12} = F\psi_y - G\psi_x$$

$$L_{21} = F'\phi_x + G'\psi_x, \quad L_{22} = F'\phi_y + G'\psi_y, \quad \overline{L}_{21} = -E\phi_y + F\phi_x, \quad \overline{L}_{22} = -F\phi_y + G\phi_x$$

$\mu = \dfrac{H}{H'}$. Now $\overline{L}_{ij} = \mu L_{ij}$ gives us the relations previously derived by other methods (1).

2. In an isometric coordinate system, $a_{ij} = b_{ij} = \delta^i_j$ (Kronecker delta), so that $|A| = |B| = 1$, $\mu = J^{\frac{n-2}{n}}$, $L_{ij} = \phi_{ij}$, $\overline{L}_{ij} = \Phi_{ij}$ and

$$\Phi_{ij} = J^{\frac{n-2}{n}} \phi_{ij}$$

When $n = 2$, these become the Cauchy-Riemann equations.

Let $W = (y_1, \cdots y_n)$ be a variable vector for which $y_i = \phi_i(x_1, \cdots x_n)$. Then $\dfrac{\partial W}{\partial x_j} = W_j = (\phi_{1j}, \phi_{2j}, \cdots \phi_{nj})$. The outer product of the $n-1$ vectors $W_1, W_2, \cdots W_{j-1}, W_{j+1}, \cdots W_n$ is a vector Ω_j whose components are $(\Phi_{1j}, \Phi_{2j}, \cdots \Phi_{nj})$. When

$$\Phi_{ij} = J^{\frac{n-2}{n}} \phi_{ij}, \quad \Omega_j = J^{\frac{n-2}{n}} W_j.$$

But Ω_j is orthogonal to the vector space spanned by the $n-1$ vectors mentioned above, so that W_j is orthogonal to this space. Hence the set of vectors $W_1, W_2, \cdots W_n$ is an orthogonal set.

1. H. Randolph Pyle, Conformal Mapping of Surfaces. Duke Mathematical Journal. Vol. II, No. 2, June, 1944.

Whittier College, Whittier, California

CARRY-OVER

Colonel John W. Ault

It is probably safe to say that every mathematics teacher believes that any student will derive great benefit from a broad background in mathematics even though his life work may be in a field which requires only a basic skill in arithmetic. This feeling is based on the fact that there is a certain amount of "carry-over" of the mental discipline to which he has been exposed.

Without any attempt to open the whole subject of carry-over, since it is one that most people consider debatable, the following remarks are devoted to the simple approach to problem solving. By the simple approach, we mean the basic procedure of isolating the question to be answered and listing the given facts which can be brought to bear upon the question. It is far too evident that most students believe that (a) Given, (b) To Find, and (c) Solution approach to a problem should be left behind at the completion of high school plane geometry.

It is believed that if the student can be made to idealize this basic approach to problems, whether they be in mathematics, physics, economics or whatever, that there will be some carry-over. In a military school it is normal to require that a student's room be kept in apple-pie order at all times, but if he does not build up in his mind the fact that neatness is a better way of life and idealize this trait to the point where he is slightly dismayed by others who do not share this belief, then his first bachelor quarters after graduation are pretty apt to be sadly disorganized.

Just how the teacher can build up this idea to the ideal level is hard to say. Of course, great benefit can be derived from constant emphasis of the principle in lectures and drill sessions, and grasping every opportunity to show the student who has difficulty with a problem that he would have avoided his difficulty by using this approach. Perhaps in the long run the most effective teaching device will be the fact that the teacher himself applies this method to all problems met in mathematics or elsewhere.

United States Air Force Academy
Colorado

A NOTE ON SIMPLE CORRELATION

Maurice Fréchet

Translation from the French
by
C. de la Menardiere
Department of Modern Languages
University of Alabama

In Volume 22, No. 2, 1958, pp. 57-69 of Mathematics Magazine, appeared an article by C. D. Smith entitled, "On the Mathematics of Simple Correlation." The aim, as given in the article, is to state the problem as it usually occurs and show how the measures of correlation may be derived by elementary algebraic calculations. This aim was completely fulfilled. The purpose of this Note is to add some comment regarding interpretation of measures of correlation which follow from theoretical considerations.

The paper states, p. 58, "The method of correlation is designed to measure the strength of common influences between two variables." We may note first that the coefficient of correlation r does not always indicate what influences may exist. Yule, and others, (1), have found cases where the measure has a value near its maximum where no common influence is known to exist between the two variables under consideration.

In fact, the thing that can best measure the coefficient of correlation is the thing that can measure a good index of correlation. We seek the total accuracy of the determination of one of the two variables when the other is given. In other words, an index of correlation has nothing to say as to the causes of this accuracy or the lack of accuracy, not even the existence of the causes. It gives a numerical fact. For instance X_1 is more or less determined by knowing X_2. This is why the term, "nonsense correlation," seems deceiving to us. In all cases, the value of an index of correlation has a meaning if the numerical data from which it has been extracted have been correctly taken. But that meaning is only one of the characteristics of the numerical table of correlation.

When a good index is equal to 1, we may state four conditions, namely:

I. To all values of X_2 there corresponds only one value of X_1.

II. Conversely when it is equal to 0.

III. X_1 is constant when X_2 varies.

IV. And conversely.

Moreover it is known that the classic coefficient, r, verifies only two of

these conditions (I and IV) (1). In some cases it is a bad index, the use of which must be controlled. The value of a good index of correlation does not dictate to the statistician the above conclusion. It invites him to look for the causes of the mathematical dependence, when for instance, the index is close to one.

Consider the coefficient of linear correlation given by the mathematical formula:

$$r = \frac{\Sigma x_1 x_2}{N\sigma_1\sigma_2}. \tag{1}$$

Also consider the correlation ratio, η as defined by $\eta = \dfrac{\sigma_m}{\sigma_1}$.

(2) We ask: 1. Why is it a correlation measure?

2. Why η is introduced if r is already known?

Karl Pearson, having noticed the faults above concerning r, tried to make a new formula.

A calculation, quite elementary, permits one to see better the passing from r to η (1).

Indeed if one has $\rho = \frac{r}{\eta}$, one has according to (1) and (2)

$$\rho = \frac{\Sigma x_1 x_a}{N\sigma_m\sigma_2}.$$

Then since $\Sigma x_2 = 0$,

$$\rho = \frac{\Sigma(x_1 - m)x_2}{N\sigma_m\sigma_2}$$

where m is the average of X_1 for $x_2 = x_a$ and σ_m^2 is the variance of m. From this it follows that:

$$r = \rho\eta$$

Here ρ is simply what becomes r when one replaces the set of points X_1, X_2 by the line of average X_1, when X_2 is given. Then ρ is, like $|r|$, at the most equal to 1 in absolute value and one has $|r| \leqq \eta \leqq 1$. One sees that ρ does not vary when the set of points (X_1, X_a) stretches or expands around the line of average.

Consequently the coefficient r contains a strange factor, ρ, which tells us nothing about the proposed sum. Is the set spread or contracted? Are the values of X_1 more or less determined when one knows X_2? By eliminating that strange factor, one has this time an index η, which verifies three of the conditions (I, II, III, IV). And even if it does not verify

exactly the condition III, which is to say, if when it equals zero, X_1 is not necessarily constant, at least the average of X_1 is constant, for given X_2, when X_2 varies.

Sometimes one considers that η is much more sensitive than r to the grouping of data by intervals and some uncertainty arises from grouping. In my opinion, this fact indicates superiority of η over r. Because the grouping by intervals modifies not the real correlation between X_1 and X_2 but that which is represented on the correlation table. An intex of correlation should reflect this change and consequently warn the statistician not to much reduce nor to much spread the grouping intervals.

It would be useful to add that, in spite of the superiority of η over r, neither r nor η will satisfy the four conditions above. Other authors, pursuing the improvement realized by Karl Pearson, have defined several correlation indexes satisfying the conditions I, II, III, IV. One such measure for instance is the index called connection by Gini (2). It would be desirable, for practical purposes, that the conclusions drawn from the calculation of r and η by means of the calculation of the index of connection of Gini be controlled. The numerical calculation was considerably simplified by Salvenini.

We have explained in what manner r is inferior to η (itself insufficient) establishing by an elementary and rapid calculation the formula

$$r = \rho\eta.$$

Here ρ is the linear coefficient of correlation that one could obtain replacing all the values of Y, when X is given, by their average. It is a geometric characteristic of the curve of the averages, equal to 1 when that curve is a straight line, and, it is only in the calculation of η that the dispersion of Y intervenes when X is given.

Since the coefficient r gives satisfactory results when the couple (X_1, X_2) satisfies the Laplace-Bravais Law, called normal to two variables, one expects to go back to that case, by transformation of the scales of X_1 and X_2 separately, in order that each of the new variables satisfies the, "second Law of Laplace," called normal Law. We have shown (3) that even after this transformation, the intended result has not been reached; the new law of the couple (X_1, X_2) modified is not necessarily the binormal law. When the frequency laws, called marginal, of X_1 and X_2, considered separately, are given, the linear coefficient of correlation r of any tables of correlation having those margins remains included between two numbers r_1, r_2, which are generally not equal either to -1 nor to 1, but simply included between -1 and $+1$ (4). These values of r_1, r_2, of contrary sign, correspond to two monotonic functional relations R_1, R_2

between X_1 and X_2. And reciprocally, if a monotonic functional relation exists between X_1 and X_2, it is identical to R_1 or R_2 (5).

BIBLIOGRAPHY

Maurice Frechet (1) *About two simple relations between the coefficient of correlation and the rapport of correlation.* (Mathematics convention, Slavic Country, Prague, 1934, pp. 209, 210).

(2) *Old and new indexes of correlation. Their application to the calculus of economical delay.* (Econometrica, XV-1947, pp. 1-30).

(3) *About the tables of correlation whose margins are given.* (C. R. Acad. Sc. V. 242, 1956, p. 2426).

(4) *On an unfounded trial to justify the coefficient of correlation*, to be completed on the proofs.

(5) *About the table of correlation whose margins are given and their applications to linear programs.* (Review of the International Institute of Statistics).

Institute Poincaré
Univ. of Paris

SOME FURTHER NOTES ON THE THEORY OF CORRELATION

C. D. Smith

The paper, referred to in the Note of Fréchet, proceeds from the Regression Model as given by Yule, (1). The Note of Fréchet adds comments on the measures r and η, listing certain limitations which follow from theoretical developments. In this note we add some comments which follow from the Regression Model, (1). Some additional references are given.

1. It seems that Karl Pearson first used the Normal Probability Surface for two variables as a model, and defined the correlation r as the product moment of deviations of the variables from their respective means, with the standard deviations as units of measure. The two straight lines located on the plane of (X_1, X_2) by the method of least squares were used by Yule as the corresponding regression model for representing r, (1), page 188. The product formula for the frequency model is $r = \dfrac{\Sigma f x_1 x_2}{N\sigma_1\sigma_2}$. Yule gives the equivalent formula $r = \dfrac{\sigma_{1e}}{\sigma_1}$ from the regression equation $x_1 = \dfrac{\Sigma x_1 x_2}{\Sigma x_2^2} x_2$, where σ_{1e} is the standard deviation of the linear estimates.

2. When the points in the x_1-arrays spread widely about the corresponding points of the regression line, the value of r is relatively small, and estimates of x_1 are not materially improved by using linear regression. When points in the arrays cluster closely about the corresponding points of the regression line, the value of r is relatively large, and estimates of x_1 may be improved by using linear regression.

3. The correlation ratio η is not less than r for a set of values (X_1, X_2) because the value of η is determined by the manner in which values within an array vary about the mean of the array, and the dispersion about the mean is less than that for any other point of an array. The values of r and η are equal for a given set (X_1, X_2) if the means of arrays fall on the regression line. The values of r and η are abstract numbers calculated without regard to the manner in which values of X_1 and X_2 may be paired. Consequently they give no information regarding possible causes of correspondence. The worker must plan his problem so that some reasonable meaning may be assigned to the results.

4. Values of correlation may be relatively large in some cases where there is no apparent cause for correspondence. H. L. Rietz illustrated the extreme case where η is greater than r by using the function $y = \sin x$, $0 \leqq x \leqq 2\pi$. Here $r=0$ and $\eta = 1$. The comparison illustrates a case of non-linear regression.

5. Yule, (1), page 195, refers to correlation as a measure of the strength of the influence of one variable upon another. The measure does not identify the influence.

6. R. A. Fisher, (2), page 190, states the case as follows: "The correlation between A and B measures, on a conventional scale, the importance of the factors which (on a balance of like and unlike action) act alike in both A and B, as against the remaining factors which affect A and B independently."

The case for common factors may be illustrated by the following experiment. Toss 5 coins on the table and count the number of heads (X_1). Leave the 3 coins nearest the edge of the table and toss the other 2 coins down with them. Count the number of heads (X_2). Repeat the experiment to obtain a sample of pairs (X_1, X_2). The two series of counts are correlated and we say the 3 coins common to the two counts of heads constitute the common factors. In performing the experiment there could be other factors such as imperfect coins, or bias in the manner of tossing them.

7. Rietz, (3) pp. 80-81, gives a probability model for correlation in the form $\phi(x, y)dxdy = g(x)h(x, y)dxdy$. When $h(x, y)$ is a function of both x and y, the two variables are said to be correlated.

LIST OF REFERENCES

(1) G. Undy Yule; Introduction to the Theory of Statistics, 6th ed. 1922.

(2) R. A. Fisher; Statistical Methods for Research Workers.

(3) H. L. Rietz; Mathematical Statistics, Carus Mathematical Monographs, No. 3.

University of Alabama

MISCELLANEOUS NOTES

Edited by

Charles K. Robbins

Articles intended for this department should be sent to *Charles K. Robbins, Department of Mathematics, Purdue University, Lafayette, Ind.*

A FURTHER GENERALIZATION OF NEUSTADT'S LAW

George Bergman

In his article "A Mathematical Munchausen" in the Nov.-Dec. '56 issue of this publication, Paul Schillo presented, amidst a large quantity of satire, a sequence of numbers called "Neustadt's Base" and a certain conjecture, "Neustadt's Law", concerning this sequence. The sequence was: .2, −1, 2, −2, 1, −.2; and the "law" that if this series be multiplied, term by term by any succession of arithmetic progressions, the resulting sequence will have zero sum. E. g.:

.2	−1	2	−2	1	−.2
×1	×2	×3	×4	×5	×6
.2	−2	6	−8	5	−1.2
×5	×3	×1	×−1	×−3	×−5
1	−6	6	8	−15	6
×1.80	×1.44	×1.08	×.72	×.36	×0
1.80	−8.64	6.48	5.76	−5.40	0

$$1.80 - 8.64 + 6.48 + 5.76 - 5.40 + 0 = 0$$

In the Jan.-Feb. '58 issue, W. W. Funkenbusch stated the correction that the number of successive multiplications must be limited to 4, and made some observations concerning multiplication by progressions of higher degree.

However, all this is merely one case of a much more general law, with numerous applications, namely:

If P is a polynomial of degree less than n,

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} P(i) = 0$$

The proof of this consists of three parts:

a) A polynomial P of degree less than n can be written in the form $P(x) = a\binom{x}{n-1} + Q(x)$ where Q is a polynomial of degree less than $n-1$

To show this, let A equal the coefficient of x^{n-1} in $P(x)$ (which may be equal to zero). Now $\binom{x}{n-1}$ is a polynomial in x of degree $n-1$, in which the coefficient of x^{n-1} is $\frac{1}{(n-1)!}$. Thus, if we subtract $(n-1)!A\binom{x}{n-1}$ from $P(x)$, we will get a polynomial of degree less than $n-1$. I.e., letting $(n-1)!A = a$, we have

$$P(x) - a\binom{x}{n-1} = Q(x)$$

or

$$P(x) = a\binom{x}{n-1} + Q(x) \qquad (Q \text{ of degree } n-1)$$

QED.

b)
$$\sum_{i=0}^{n} (-1)^i \binom{n}{i}\binom{i}{n-1} = 0$$

The only values of i for which $\binom{i}{n-1} \neq 0$ within the indicated range are n and $n-1$.

Thus

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i}\binom{i}{n-1} = (-1)^n \binom{n}{n}\binom{n}{n-1} + (-1)^{n-1}\binom{n}{n-1}\binom{n-1}{n-1}$$

$$= (-1)^n \cdot 1 \cdot \binom{n}{n-1} - (-1)^n \binom{n}{n-1} \cdot 1 = 0$$

c) If for all polynomials $Q(x)$ of degree less than $n-1$,

$$\sum_{i=0}^{n-1} (-1)^i \binom{n-1}{i} Q(i) = 0,$$

then for all polynomials $P(x)$ of degree less than n,

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} P(i) = 0.$$

To show this we begin by applying a, and then b to the expression given:

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} P(i) = \sum_{i=0}^{n} (-1)^i \binom{n}{i}\left[a\binom{i}{n-1} + Q(i)\right]$$

$$= a\sum_{i=0}^{n}(-1)^i\binom{n}{i}\binom{i}{n-1}+\sum_{i=0}^{n}(-1)^i\binom{n}{i}Q(i)=\sum_{i=0}^{n}(-1)^i\binom{n}{i}Q(i).$$

Now, as Q is a polynomial of degree less than $n-1$, we can prove this last expression equal to zero if we can break it into expressions of the form given in our hypothesis. This can be done as follows:

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}Q(i)=\sum_{i=0}^{n}(-1)^i\left[\binom{n-1}{i}+\binom{n-1}{i-1}\right]Q(i)$$

$$=\sum_{i=0}^{n}(-1)^i\binom{n-1}{i}Q(i)+\sum_{i=0}^{n}(-1)^i\binom{n-1}{i-1}Q(i).$$

The final term of the first of these sums is zero, and first term of the second sum is zero. We can adjust the ranges accordingly. To make the second conform with the form in the hypothesis, we shall then substitute $j=i-1$. $Q(j+1)$ can be expressed as a simple polynomial in j, of the same degree as Q, which we shall call Q':

$$\sum_{i=0}^{n}(-1)^i\binom{n-1}{i}Q(i)+\sum_{i=0}^{n}(-1)^i\binom{n-1}{i-1}Q(i)$$

$$=\sum_{i=0}^{n-1}(-1)^i\binom{n-1}{i}Q(i)+\sum_{i=1}^{n}(-1)^i\binom{n-1}{i-1}Q(i)$$

$$=0+\sum_{j=0}^{n-1}(-1)^{j+1}\binom{n-1}{j}Q(j+1)$$

$$=0-\sum_{j=0}^{n-1}(-1)^j\binom{n-1}{j}Q'(j)=0-0=0$$

This proves c), which, together with the self-evident case when $n=0$, constitutes an inductive proof of the theorem.

Now "Neustadt's Base" merely consists of the numbers $(-1)^i\binom{5}{i}$, disguised by division by 5. Since the product of less than five arithmetic progressions is a progression whose general term is a polynomial of degree less than five, the sum of the products of corresponding terms of

less than five arithmetic progressions, multiplied by the corresponding terms of Neustadt's Base, will equal zero divided by five, or zero. ("I. Neustadt" did not believe in mathematical induction – but one can't please everybody! The particular case of "Neustadt's Base", or any particular case, can be proved deductively by multiplying the numbers of this "base" successively by a, $a+d$, $a+2d$, $\cdots$; then by a', $a'+d'$, $\cdots$; then a'', $a''+d''$, $\cdots$.)

The applications, as I mentioned, are numerous. For instance, take the method commonly used for finding the degree of a progression $\{S_n\}$. We find first the differences between successive terms, ΔS, and then the differences of the differences, $\Delta^2 S$, and so forth, until we reach an n such that all the $\Delta^n S$ equal zero; whereupon we can state that $n-1$ is the degree of the progression. This can be explained in terms of the above theorem when one considers the fact that the kth $\Delta^n S$ is

$$\sum_{i=0}^{n} (-1)^{n+i}\binom{n}{i} S_{i+k}\,.$$

Another use for it is in the summing of progressions whose general term is $P(x)r^x$.* Since the numbers $(-1)^i\binom{n}{i}$ are the coefficients of the r^i in the expansion of $(1-r)^i$, if one multiplies the series by $(1-r)^{d+1}$ where d is the degree of $P(\ )$, then the coefficient of r^x becomes

$$\sum_{i=0}^{d+1} (-1)^i \binom{d+1}{i} P(x-i)\,,$$

or 0, except for the first and last d terms, where the sum will be incomplete for want of non-included terms. Thus, all but $2d$ terms will drop out, and dividing through by $(1-r)^i$, one will have a closed expression for the sum.

Example : Evaluate

$$S = 1 + 4r + 9r^2 + \cdots + (n-2)^2 r^{n-3} + (n-1)^2 r^{n-2} + n^2 r^{n-1}$$

Solution : Multiply through by $1-3r+3r^2-r^3$ to obtain

$$(1-r)^3 S = 1 - r + 0 + \cdots + 0 + (-n^2-2n-1)r^n + (2n^2+2n-1)r^{n+1} - n^2 r^{n+2}\,.$$

Thus
$$S = \frac{1-r-(n^2+2n+1)r^n+(2n^2+2n-1)r^{n+1}-n^2 r^{n+2}}{(1-r)^3}.$$

*as, for instance, Proposal 333. $r = \frac{x-y}{y}$

Stuyvesant High School, New York, N. Y.

A NEW LOOK AT AN OLD PROBLEM

Charles E. Wingo Jr.

I have always been deeply interested in mathematics, and for the last ten years mathematics has been a source of recreation and entertainment for me. It was thru this that the old problem

$$x^2 + y = 11, \quad y^2 + x = 7$$

was revived. It was over fifty years ago that I first heard of it. What has stimulated interest in this problem is the fact that it has proved to be as much a puzzle as a problem.

Equations of the form

$$x^2 + y = a, \quad y^2 + x = b$$

were pronounced insoluble by Prof. Quimby of Dartmouth, except by Descartes formula for solving bi-quadratics.

Those who are further interested in some of the solutions to the problem

$$x^2 + y = 11, \quad y^2 + x = 7$$

are refered to the Am. Mathematical Monthly, Vol. V, pages 291-2-3-4-5 and Vol. VI, pages 13-37, where will be found about eight solutions, some of which are quite complicated; bi-quadratics, factoring, and other methods.

In the problem as treated here, an effort is made to develope a method of solution whereby all problems of the form

$$x^2 + y = a, \quad y^2 + x = b$$

can be easily solved, where $x > y$ and x, y, a and b are positive integers, provided of course that integral solutions exist. This means that arbitrary values of a and b can not be taken, as

$$x^2 + y = 12, \quad y^2 + x = 4.$$

The method derived here is general in scope, subject to the limitation mentioned above. Consider the two equations

$$x^2 + y = a$$

and

$$y^2 + x = b.$$

We have

$$x^2 - y^2 - (x - y) = a - b$$

or

(1) $$(x-y)(x+y-1) = a-b$$

In my study of this problem over the past few years, my interest centered on equation (1) above. For the solution, if any was to be found, would hinge on this equation. For it is obvious that $(a-b)$ is the product of two factors $(x-y)$ and $(x+y-1)$, and the key to the solution was to find which of these equate to $(x-y)$ and which to $(x+y-1)$.

At this stage the problem rested for some time. Then some time later it flashed to me that $\sqrt{x^2+y}$ and $\sqrt{y^2+x}$ could be written

$$\sqrt{x^2+y} = \sqrt{a} = x + c$$

$$\sqrt{y^2+x} = \sqrt{b} = y + d$$

$$\sqrt{a} - \sqrt{b} = x - y + c - d$$

Since $x > y$, we know that in taking the square roots of $x^2 + y$ and $y^2 + x$, the amount that c adds to x is less than the amount that d adds to y. Therefore d is larger than c. So $\sqrt{a} - \sqrt{b}$ will be less than the true value of $x - y$ by the amount of $d - c$. So the value of the factor of $a - b$ to equate to $x - y$ will be the integral value next greater than $\sqrt{a} - \sqrt{b}$ and $(x + y - 1)$ would be equal to the other factor.

To show how the method is applied to problems we consider the following.

$$x^2 + y = 11 = a$$

$$y^2 + x = 7 = b$$

$$a - b = 4$$

Factors of $4 = 1 \cdot 2 \cdot 2$

$$\sqrt{a} = \sqrt{11} = 3.31$$

$$\sqrt{b} = \sqrt{7} = 2.67$$

$$\sqrt{a} - \sqrt{b} = .65$$

The first integral factor larger than .65 is 1. So

$$x - y = 1$$

(continued on page 288)

PROBLEMS AND QUESTIONS

Edited by

Robert E. Horton

Readers of this department are invited to submit for solution problems believed to be new and subject matter questions that may arise in study, in research, or in extra-academic situations. Proposals should be accompanied by solutions, when available, and by any information that will assist the editor. Ordinarily, problems in well-known textbooks should not be submitted.

Solutions should be submitted on separate, signed sheets. Figures should be drawn in India ink and twice the size desired for reproduction.

Send all communications for this department to *Robert E. Horton, Los Angeles City College, 855 North Vermont Ave., Los Angeles 29, California*

PROPOSALS

376. *Proposed by C.W.Trigg, Los Angeles City College.*

Identify the following unorthodox definitions: (1) the earthy or stony substance in which an ore or other mineral is bedded, (2) fully sufficient, (3) the four brightest components of θ Orionis, (4) type of speech fancifully exaggerated, (5) oddity, (6) a Franciscan friar, (7) a deposit formed in a liquid vegetable extract, (8) a large molding of convex profile, (9) fancied, (10) the series of air bubbles made by the breath from an otter under water, (11) a drain to carry off filthy water, (12) one sixteenth of a fluid drachm, (13) scheme, (14) depicted as walking, (15) absent-minded, (16) seekers, (17) complete, (18) invalid, (19) an interpreter of music.

377. *Proposed by J.M.Gandhi, Thapar Polytechnic, Patiala, India.*

Prove that

$$\sum_{i=0}^{3n+1} \binom{6n+2}{2i} 3^{i} \equiv 0,\ 2^{3n+1},\ -2^{3n+1}, \quad [\bmod\ 2^{3n+2}]$$

when n is of the form $2m$, $4m+3$ or $4m-1$ respectively.

378. *Proposed by Barney Bissinger, Lebanon Valley College, Pennsylvania.*

Let a_0, a_1 be arbitrary and $n(n+1)a_{n+1} = n(n-1)a_n - (n-2)a_{n-1}$ for $n > 0$. Find $S = \sum\limits_0^{\infty} a_n$.

379. *Proposed by J.M.Howell, Los Angeles City College.*

Two men fight a duel. They both fire at a given signal. If both are

alive, they fire again at a given signal, and repeat the process until at least one of them is dead. If the probability that A kills B on any round is p, and the probability that B kills A on any round is r, find the probability that A is alive, B is alive or neither is alive after: 1) n rounds, 2) after an infinite number of rounds.

380. *Proposed by Huseyin Demir, Kandilli, Eregli, Kdz., Turkey.*

Solve the system of equations

$$x(z-a)+u(x+u) = 0$$

$$y(x-b)+u(y+u) = 0$$

$$z(y-c)+u(z+u) = 0$$

where $abc \neq 0$ and $a^{-1}+b^{-1}+c^{-1} = u^{-1}$.

381. *Proposed by George M. Bergman and Melvin Hochster, Stuyvesant High School, New York.*

Show that if the circumsphere and the insphere of a tetrahedron are concentric, four spheres can be drawn each tangent to each edge of the tetrahedron, extended if necessary.

382. *Proposed by C.N.Mills, Sioux Falls College, South Dakota.*

For the hypocycloid $x^{2/3}+y^{2/3} = a^{2/3}$, (α, β) is the center of curvature. Show that $\alpha+\beta = (x^{1/3}+y^{1/3})^3$.

SOLUTIONS

Late Solutions

350. *J.L.Brown, Jr., Ordnance Research Laboratory, Pennsylvania State University; M.Morduchow, Polytechnic Institute of Brooklyn.*

The Wine Merchant

355. [November 1958] *Proposed by P.B.Jordain, New York, New York.*

A wine merchant had a small cask containing 100 gallons of wine. In order to make more money, he decided to replace each gallon he took out of the cask by a gallon of water. This he did n times. Finding that he was losing customers, he naively tried to undo what he had done by selling his watered wine from the same cask, but replacing each gallon sold of the mixture by a gallon of pure wine, feeling that by this method, at the end of an additional n gallons of watered wine sold, he would have a cask of full wine again. He was unfortunate enough in having hit upon a total number of gallons sold such that the wine contained in the cask

was at a minimum at the end of this operation. The question is, how many gallons of watered wine had he sold in all? Assume that the merchant dealt only in full gallon sales.

Solution by David F. Finlayson, Worcester Polytechnic Institute, Massachusetts. After n sales the fraction of wine remaining would be given by $(\frac{99}{100})^n$. The gallons of wine remaining would be given by $100(\frac{99}{100})^n = \frac{(99)^n}{(100)^{n-1}}$. The gallons of water left would be given by $100 - \frac{(99)^n}{(100)^{n-1}}$. After repeating this process n more times by replacing the mixture sold by pure wine instead of water as in the first n sales, the gallons of water left would be given by $(\frac{99}{100})^n[100 - \frac{(99)^n}{(100)^{n-1}}]$. The wine in the cask would then be given by

$$100 - (\frac{99}{100})^n[100 - \frac{(99)^n}{(100)^{n-1}}] = G.$$

or

$$G = 100 - \frac{(99)^n}{(100)^{n-1}} + \frac{(99)^{2n}}{(100)^{2n-1}}$$

$$G = 100 - (99)^n(100)^{1-n} + (99)^{2n}(100)^{1-2n}$$

Differentiating with respect to n and setting $\frac{dG}{dn} = 0$ gives

$$\frac{dG}{dn} = -\log_e 99 + \log_e 100 + (99)^n \log_e 99(100)^{-n}(2) + (100)^{-n}\log_e 100(99)^n(-2) = 0$$

$$\log_e 99 - \log_e 100 = 2(\frac{99}{100})^n(\log_e 99 - \log_e 100)$$

$$1 = 2(\frac{99}{100})^n$$

$$0.5 = (0.99)^n$$

$$n = 69\ *$$

$$2n = 138 \qquad 2n - 1 = 137$$

Since the first gallon sold was pure wine, the wine merchant sold $2n - 1 = 137$ gallons of watered wine.

Substituting the value of n in the expression for the number of gallons of wine left, we find that the wine merchant ends with 83.57 gallons of wine in the cask.

*It can be immediately seen that $(.99)^{69}$ can not equal exactly 0.050 since the last digit of .99 to any power can not be 0. Solving the equation by means of a slide rule or logarithms produces an answer very close to 69. It is better, however, to square both sides of the equation and solve for $2n$. By doing this an irrational answer between two succesive integers will be found and of course the even integer is the answer desired. The value for $2n$ is then found to be 138 and n must equal 69 and not 68 or 70.

Also solved by Stephen A. Andrea, Oberlin, Ohio; Philip Fung, Idaho State College; Joseph D. E. Konhauser, Haller Raymond and Brown, Inc., State College, Pennsylvania; Sam Kravitz, East Cleveland, Ohio; Richard Mittleman, Los Angeles City College; Lawrence A. Ringenberg, Eastern Illinois University; C. W. Trigg, Los Angeles City College; Dale Woods, Idaho State College; and the proposer.

Several solutions were received with the value of n incorrectly approximated by 68 instead of 69. Ringenberg pointed out that this problem is a special case of problem E962 appearing in the March-1952 issue of the AMERICAN MATHEMATICAL MONTHLY.

A Reciprocal Sum

357. [November 1958] *Proposed by Joseph Andrushkiw, Seton Hall University, New Jersey.*

Show that if $d>0$,

$$\sum_{k=0}^{\infty} 1/(k^2-2dk+2d^2) = k+3\pi/4d, \quad 0 \leq k < 1/2d^2$$

Solution by Arne Pleijel, Trolhattan, Sweden. Let the proposal be restated in the form:

$$0 \leq \sum_{k=0}^{\infty} \frac{1}{k^2-2dk+2d^2} - \frac{3\pi}{4d} \leq \frac{1}{2d^2}.$$

Then we have

$$\frac{3\pi}{4d} = \int_0^{\infty} \frac{dx}{x^2-2dx+2d^2}$$

$$= \int_0^{\infty} \frac{dx}{(x-d)^2+d^2}$$

$$= \sum_{k=0}^{\infty} \int_k^{k+1} \frac{dx}{(x-d)^2+d^2} \; .$$

and thus, as $\frac{1}{(x-d)^2+d^2}$ is decreasing in x, we have

$$\sum_{k=0}^{\infty} \frac{1}{(k+1-d)^2+d^2} \leqq \frac{3\pi}{4d} \leqq \sum_{k=0}^{\infty} \frac{1}{(k-d)^2+d^2}$$

or

$$-\frac{1}{2d^2} + \sum_{k=0}^{\infty} \frac{1}{(k-d)^2+d^2} \leqq \frac{3\pi}{4d}$$

and so the proposition is proved.

Also solved by the proposer.

The Bouncing Ball

358. [November 1958] *Proposed by C. W. Trigg, Los Angeles City College.*

A ball having fallen from rest a vertical distance h, strikes a stone protruding from a wall and bounces off horizontally without spinning. If the distance of the stone from the ground is s and the coefficient of restitution is e, show that

a) the ball will strike the ground at a distance $2\sqrt{she}$ from the foot of the wall;

b) the inclination of the stone to the horizontal is $\arctan\sqrt{e}$.

Solution by Philip Fung, Idaho State College. Let u, v be the velocities of ball before and after impact and their components parallel and perpendicular to the surface of stone be u_x, u_y and v_x, v_y respectively. Let the inclination of stone to the horizontal be θ, the time taken by the ball to strike the ground after impact with stone t, and the horizontal distance covered by the ball during this time be R.

Now, the velocity of ball just before impact with the stone is $u = \sqrt{2gh}$. So

$$u_x = u \operatorname{Sin} \theta;$$

$$u_y = u \operatorname{Cos} \theta.$$

Since u_x is parallel to the stone surface and thus unaffected in magnitude

and direction by impact, therefore we have,

$$v_x = u_x = u \operatorname{Sin} \theta;$$

$$v_y = eu_y = eu \operatorname{Cos} \theta.$$

But, the direction of v is horizontal. Thus,

$$\operatorname{Tan} \theta = v_y/v_x = e/\tan \theta.$$

From this we have,

$$\theta = \operatorname{arc} \tan \sqrt{e}.$$

Now,

$$s = 1/2\ gt^2$$

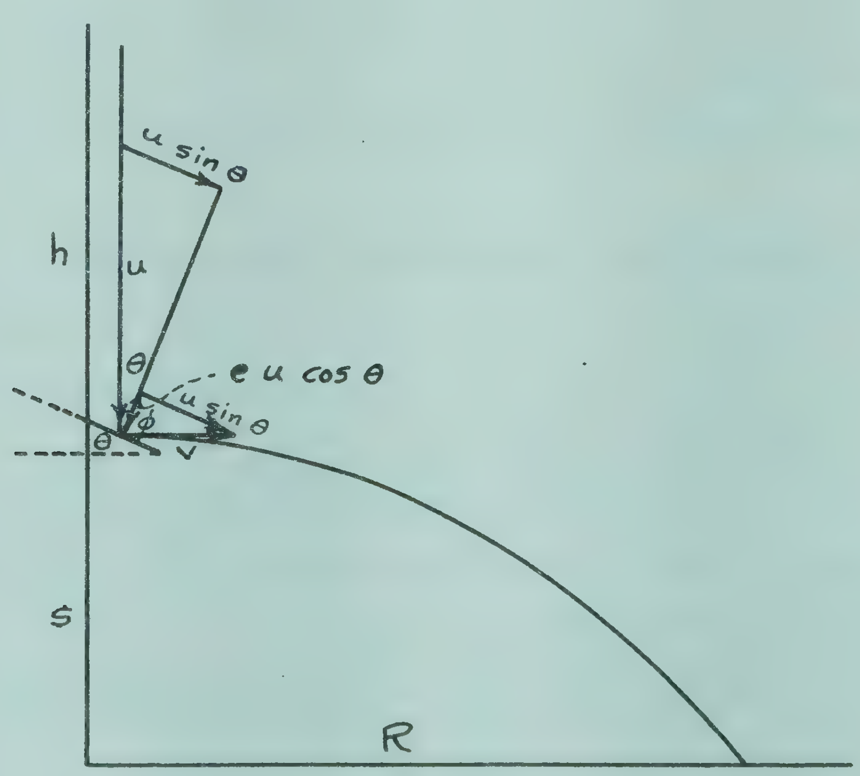
or,

$$t = \sqrt{2s/g}$$

And,

$$R = vt = t\sqrt{v_x^2 + v_y^2}$$

Hence, by substituting the values of t, v_x, v_y, and then θ into the equation above, we can thus solve that $R = 2\sqrt{she}$.



Also solved by William E. F. Apphun, St. John's University, New York; Joseph M. C. Hamilton, Los Angeles City College; Joseph D. E. Konhauser, Haller, Raymond and Brown, Inc., State College, Pennsylvania; Dale Woods, Idaho State College; and the proposer.

The String-of-Pearls Polynomial

359. [November 1958] *Proposed by Norman Anning, Alhambra, California.*

Prove that the string-of-pearls polynomial $x^k + 1$ can be expressed in at least one non-trivial way as the sum of two squares, if k is any even positive integer different from 2, 4, 8, 16 ⋯.

Solution by William M. Sanders, Mississippi Southern College. Let k be the n-th term of an arithmetic progression with first term 0; i.e., $k = (n-1)d$. Impose the condition that d be the maximum divisor of k such that $n = 2p \geqq 4$ to obtain the progression $0, d, 2d, \cdots, (n-2)d, k$. This is possible since k is even and different from 2, 4, 8, 16 ⋯. Now for $k = 2m$, $d = 2q$, and $n = 2p$,

$$x^k + 1 = x^{2m} + \sum_{v=1}^{2p-2} (-1)^v 4 \min(v, 2p-v-1)\, x^{2m-2vq}$$

$$- \sum_{v=1}^{2p-2} (-1)^v 4 \min(v, 2p-v-1)\, x^{2m-2vq} + 1$$

$$= \left(x^m + 2 \sum_{v=1}^{p-1} (-1)^v x^{m-2vq} \right)^2 + \left(2 \sum_{v=1}^{p-1} (-1)^{v+1} x^{m+q-2vq} + 1 \right)^2$$

Also solved by L. Carlitz, Duke University; Arne Pleijel, Trolhattan, Sweden; and the proposer.

A Difference Equation

360. [November 1958] *Proposed by Chih-yi Wang, University of Minnesota.*

Considering the higher differences of $\binom{n}{x}$ with respect to x show that

$$\Delta^m \binom{n}{x} = \frac{1}{(n-x)\,!} \{ D_t^{n-x} [t^n (t-2)^m] \}_{t=1}.$$

Solution by Arthur E. Danese, Union College, Schenectady, New York. In Rodrigues' formula for Jacobi polynomials [Szegö, ORTHOGONAL POLYNOMIALS, New York, 1939, p. 66, 4.3.1]:

$$(1-x)^\alpha (1+x)^\beta P_n^{(\alpha, \beta)}(x) = \frac{(-1)^n}{2^n n\,!} \left(\frac{d}{dx}\right)^n \{ (1-x)^{n+\alpha} (1+x)^{n+\beta} \},$$

replace x by $t-1$, n by $n-x$, β by x, α by $x-n+m$, and evaluate at $t = 1$ to obtain

$$\frac{1}{(n-x)!}\{D_t^{n-x}[t^n(t-2)^m]\}_{t=1} = 2^{n-x}(-1)^{x-n+m}P_{n-x}^{(x-n+m,\, x)}(0)$$

which equals

$$2^{n-x}(-1)^m P_{n-x}^{(x,\, x-n+m)}(0)$$

since

$$P_v^{(\alpha,\beta)}(u) = (-1)^v P_v^{(\beta,\alpha)}(-u).$$

The use of the explicit evaluation of the Jacobi polynomial:

$$P_n^{(\alpha,\beta)}(x) = \sum_{v=0}^{n} \binom{n+\alpha}{n-v}\binom{n+\beta}{v}\left(\frac{x-1}{2}\right)^v\left(\frac{x+1}{2}\right)^{n-v}$$

[ibid, p. 67, 4.3.2] yields

$$2^{n-x}(-1)^m P_{n-x}^{(x,\, x-n+m)}(0) = \sum_{v=0}^{n-x} (-1)^{v+m}\binom{m}{v}\binom{n}{x+v}$$

which is equivalent to $\Delta^m\binom{n}{x}$.

Also solved by the proposer.

A Variable Nine-Point Circle

361. [November 1958] *Proposed by N. A. Court, University of Oklahoma.*

A variable triangle inscribed in a rectangular hyperbola has a fixed vertex and the opposite side moves parallel to itself. Show that its variable nine-point circle passes through two fixed points.

Solution by J. W. Clawson, Collegeville, Pennsylvania. It is well known that the orthocenter of a triangle inscribed in a rectangular hyperbola lies on the hyperbola. Also that the center of the hyperbola lies on the nine-point circle. (C. Smith, Geometrical Conics, pages 165, 166.)

Let A be the fixed vertex, BC the variable side which moves parallel to itself. Let the perpendicular from A to BC intersect the hyperbola at H. Then the fixed point H is the ortho-center of the variable triangle ABC. Since the nine-point circle of a triangle passes through the mid-point of AH, the variable nine-point circle passes through this point and also through the center of the hyperbola.

Also solved by Huseyin Demir, Kandilli, Eregli, Kdz., Turkey; Sister M. Stephanie, Georgian Court College, New Jersey; and the proposer.

Comment on Problem 337.

337. [March 1958 and November 1958] *Proposed by Victor Thebault, Tennie, Sarthe, France.*

Comment by William E. F. Appuhn, St. John's University, New York.

In order to eliminate having sides less than or equal to zero, conditions should have been placed as the solution was developed:

1. In equation (1) $x = 2pq$, the condition $pq > 0$.
2. In equation (2) $y = p^2 - q^2$, the condition $p^2 > q^2$.
3. In the equations for p and q we do have to place restrictions on the integers g and h. They are:
   (a) $g \neq 0$, (b) $h \neq 0$, (c) $g^2 \neq h^2$, (d) $gh > 0$ if $g^2 > h^2$, and (e) $gh < 0$ if $g^2 < h^2$.
4. In the final set of equations, we have to place the restriction $gh > 0$.
   The above conditions will be satisfied by $gh > 0$ and $g^2 > h^2$.
5. Noting that, by equations (1) and (3), x and z are both even hence y is also even, a much quicker and easier complete solution may be obtained as follows:

Let (1) $a = 4uv$, $b = 2|u^2 - v^2|$ and hypotenuse $e = 2(u^2 + v^2) = 2d^2$ where u and v are different positive integers. In order that $u^2 + v^2 = d^2$ let (2) $u = 2xy$, $v = |x^2 - y^2|$, $\therefore d = x^2 + y^2$, where $x \neq y$ are integers. Substituting (2) in (1) to obtain the immediate solution:
(3) $a = 8xy\,|x^2 - y^2|$, $b = 2|6x^2y^2 - x^4 - y^4|$ and $c = 2(x^2 + y^2)^2$.

QUICKIES

From time to time this department will publish problems which may be solved by laborious methods, but which with the proper insight may be disposed of with dispatch. Readers are urged to submit their favorite problems of this type, together with the elegant solution and the source, if known.

Q 247. For what values of x is $m^2 + n^2 - a^2 - b^2 > (mn - ab)x$ where $0 \leqq a \leqq m$ and $0 \leqq b \leqq n$. [*Submitted by M. S. Klamkin*]

Q 248. Show that if k is any real number, the lines $x^4 + kx^3y - 6x^2y^2 - kxy^3 + y^4 = 0$ cut $x^2 + y^2 = 1$ into eight equal parts. [*Submitted by Norman Anning*]

Q 249. Show that for the smallest values of the angles, $2 \operatorname{arc} \cos \sqrt{1/3} = \operatorname{arc} \cos (1/3) = \Pi$. [*Submitted by C. W. Trigg*]

Q 250. Prove that 19^{19} can not be represented as the sum of a fourth power and a cube. [*Submitted by D. L. Silverman*]

Q 251. In a rectangular coordinate system (x, y), $y = f(x)$. If the coordinates are transformed into new ones, (x', y'), by a rotation through a positive angle α, express $\frac{dy'}{dx'}$ and $\frac{d^2y'}{dx'^2}$ in terms of $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$. [*Submitted by Henry E. Fettis*]

Q 252. Find the sum of

$$S = \binom{m}{r}\binom{n}{0} + \binom{m}{r-1}\binom{n}{1} + \cdots + \binom{m}{0}\binom{n}{r}$$

[*Submitted by M. S. Klamkin*]

Answers

A 248. The figure is invariant under a rotation through 45°.

A249. In general $(\sqrt{a/b})^2 + (\sqrt{[1+(b-2a)/b]/2})^2 = 1$. So $\arccos\sqrt{a/b} + \arccos\sqrt{[1+(b-2a/b]/2} = \pi/2$ or $2\arccos\sqrt{a/b} + \arccos(b-2a)/b = \pi$. In the particular case, $a = 1$ and $b = 3$.

A 250. $x^3 \equiv 0, 1, 5, 8$, or $12 \pmod{13}$ and $y^4 \equiv 0, 1, 3$, or $9 \pmod{13}$. Thus $x^3 + y^4$ may be congruent to anything except $7 \pmod{13}$. But $19^{19} \equiv 6^{19} \equiv 6^{12}\cdot 6^7 \equiv 6^7 \equiv 7 \pmod{13}$.

A 251. If the line tangent to $y = f(x)$ makes an angle θ with the x axis and angle θ' with the x'axis, then $\theta = \theta' + \alpha$. So

$$\frac{dy'}{dx'} = \tan\theta' = \frac{\tan\theta\tan\alpha}{1+\tan\theta\tan\alpha} = \frac{dy/dx\ \tan\alpha}{1+(dy/dx)\tan\alpha}$$

Further, since the curvature is invariant under the rotation of axes, we have

$$\frac{\frac{d^2y}{dx^2}}{[1+(\frac{dy}{dx})^2]^{3/2}} = \frac{\frac{d^2y'}{dx'^2}}{[1+(\frac{dy'}{dx'})^2]^{3/2}}$$

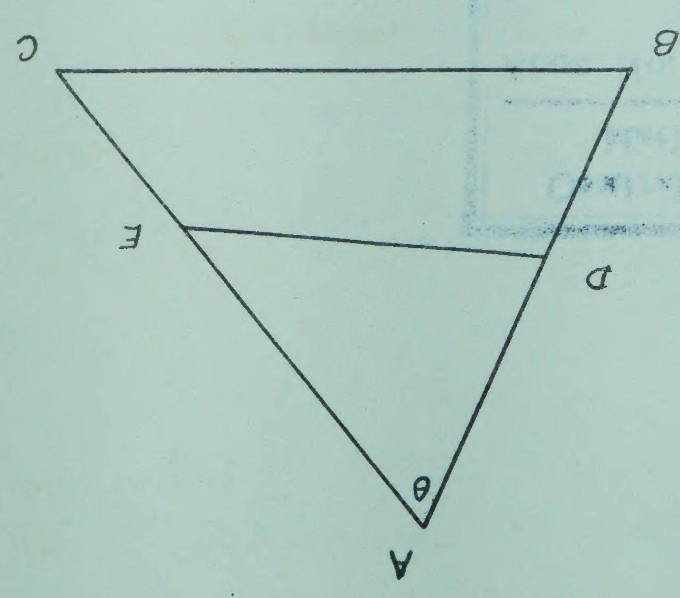
from which $\frac{d^2y'}{dx'^2}$ can easily be expressed in terms of $\frac{d^2y}{dx^2}$ and $\frac{dy}{dx}$.

A 252. On equating coefficients of x^r on both sides of the identity

$$(1+x)^m(1+x)^n = (1+x)^{m+n}$$

we have the sum $S = \binom{m+n}{r}$.

A 247. Let $x = 2\cos\theta$. Then we have $m^2 + n^2 - 2mn\cos\theta > a^2 + b^2 - 2ab\cos\theta$. In the figure, $AB = m$, $AC = n$, $AD = a$ and $AE = b$. So $\overline{BC}^2 = m^2 + n^2 - 2mn\cos\theta$ and $\overline{DE}^2 = a^2 + b^2 - 2ab\cos\theta$. In order that DE be less than BC we must have $\theta > 60°$ or $x < 1$.



(A New Look at an Old Problem – *Continued from page 276.*)

$$x+y-1 = 4$$

$$2x = 6 \quad x = 3$$

$$2y = 4 \quad y = 2$$

Next take:

$$x^2+y = 1052 = a$$

$$y^2+x = 816 = b$$

$$a-b = 236$$

Factors of 236 are $2\cdot 2\cdot 59 = 4\times 59$.

$$\sqrt{a} = \sqrt{1052} = 32.43$$

$$\sqrt{b} = \sqrt{816} = 28.55$$

$$\sqrt{a}-\sqrt{b} = 3.88$$

The integral factor which is next and larger than 3.88 is 4. Then

$$x-y = 4$$

$$x+y-1 = 59$$

$$2x = 64 \quad x = 32$$

$$2y = 56 \quad y = 28$$

For a further consideration of the problem, take the case where $a = b$. Then:

$$x^2+y = a$$

$$y^2+x = a$$

$$x^2-y^2-(x-y) = 0$$

$$(x-y)(x+y-1) = 0$$

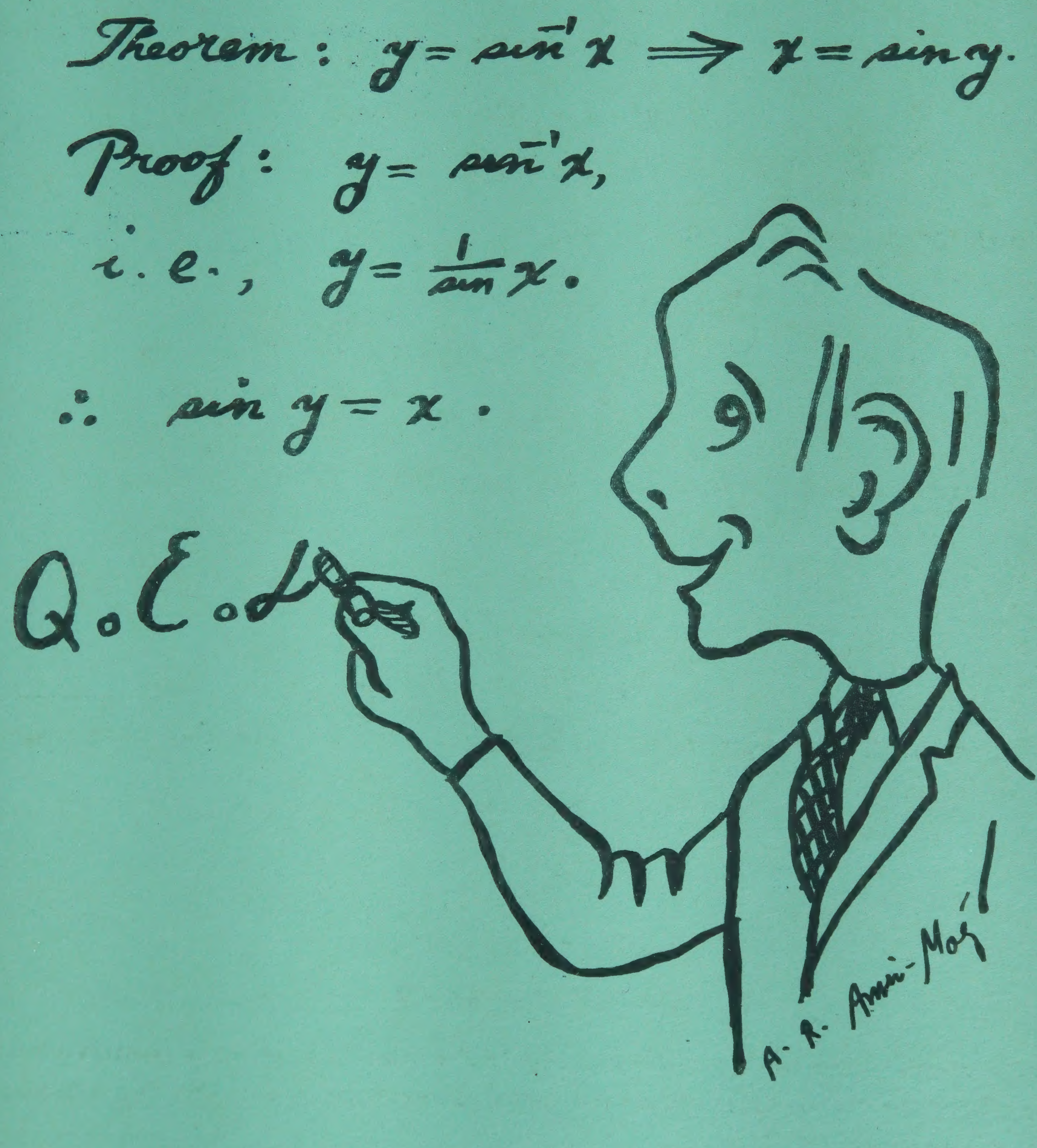
Since x and y are both positive integers, $(x+y-1)$ is also a positive integer. Therefore $(x-y)$ is equal to zero. Therefore $x = y$. So the equation above becomes:

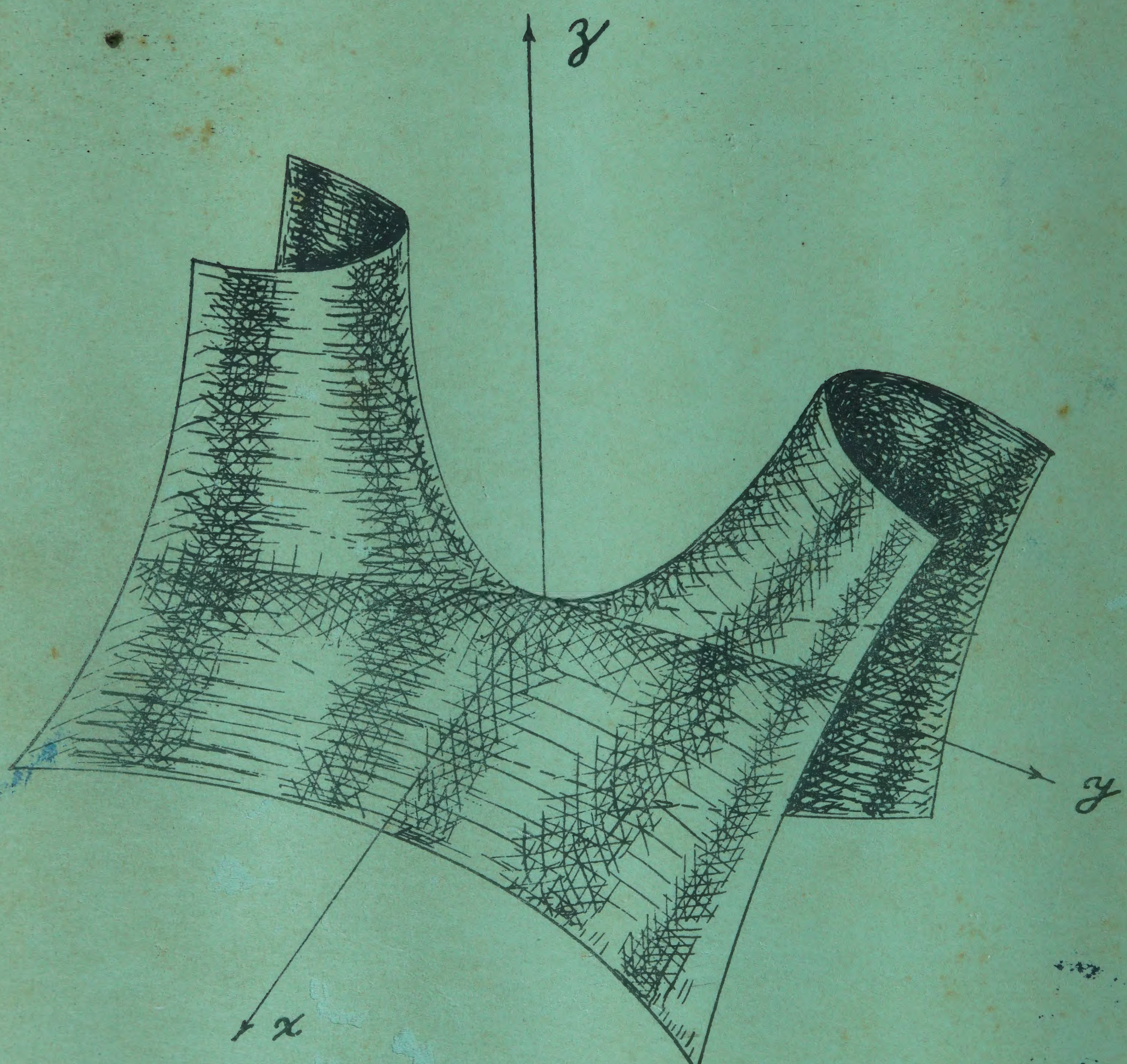
$$x^2+x = a$$

$$x(x+1) = a$$

It is seen from the equation above that a is the product of two consecutive integers. Hence a has such values as 1×2, 2×3, 3×4 etc.

Amelia, Va.


Theorem: $y = \sin^{-1} x \Rightarrow x = \sin y$.
Proof: $y = \sin^{-1} x$,
i.e., $y = \frac{1}{\sin} x$.
$\therefore \sin y = x$.
Q.E.D.
A. R. Amir-Moéz



$$\frac{x^2}{a^2} - \frac{y^2}{b^2} + 2cz = 0$$